Universidad de Oviedo

Programa de Do ctorado

en Matemáticas yEstadística

Comparison of alternatives und e uncertainty and imprecision

Tesis Do ctoral

Ignacio Montes Gutiérr ez

Conte nts

Agradecimientos

Esta memoria abarca el trabajo realizado desde que en Julio de 2009 finalicé la licenciatura en Matemáticas. Por aquel entonces recibí el apoyo de varios miembros del Departamento de Estadística e Investigación Operativa y Didáctica de la Matemática para realizar una tesis doctoral. De entre ellos, debo destacar los consejos que siempre me dieron los Profesores Pedro Gil y Santos Domínguez.

La oportunidad de comenzar este trabajo me surgió cuando Susana Montes se puso en contacto conmigo para ofrecerme la posibilidad de adentrarme en el apasionante mundo de la investigación, y es de justicia calificar este momento como uno de los más importantes de toda mi vida. Desde entonces me encuentro en el mejor ambiente posible de trabajo, arropado en todo momento por los miembros del proyecto UNIMODE: Susana Díaz, Davide Martinetti, Enrique Miranda y la propia Susana Montes.

Es de justicia nombrar de manera especial a mis directores Susana Montes y Enrique Miranda (y aunque no figure como directora, también a Susana Díaz). Este trabajo ha sido posible gracias a ellos, a su dedicación, a su paciencia, a su pasión por las matemáticas, a sus ganas de hacer las cosas bien. Podría rellenar muchas páginas explicando lo agradecido que les estoy y lo afortunado que me siento por pertenecer a este maravilloso grupo de trabajo, pero sinceramente mis palabras nunca llegarían a reflejar mis sentimientos.

A lo largo de estos años ha habido muchas personas a las que les debo mis agradecimientos, comenzando por mis compañeros de carrera y los miembros del Departamento de Estadística e Investigación Operativa y Didáctica de la Matemática. Al personal de la UCE, y en particular a Tania y Patri, por su incondicional apoyo y su sincera amistad. A Gert de Cooman y todo su equipo, y a Paolo Vicig y Renato Pelessoni debo agradecerles su amabilidad durante las estancias que realicé en las Universidades de Gante y de Trieste. A los miembros de la comunidad científica nacional e internacional por su amabilidad y buenos consejos. Y en definitiva, a todos aquellos que de una u otra manera han colaborado o aportado en mi formación.

Por otra parte, este trabajo no habría sido posible sin la financiación, a través de una

Beca deFormación deProfesorado Universitario, del Ministerio deEducación, así comoa la Ayuda recibida por parte del Vicerrectorado de Investigación y Campus de Excelencia de la Universidad de Oviedo que me p ermitió realizar una estancia de investigación en la Univers idad de Trieste.

Y por supuesto, darles lasgraciasa mi familia y a Almudena .. . simplemente por to do.

Resumen

En much as situaciones de la vida real es necesario comparar alternativas. Además , es habitual que estas alternativas estén definidas ba jo falta de información. En esta memoria se consideran dos tip os de falta de información: incertidumbre e imprecisión. La incertidumbre se refiere a situaciones en las cuales los posibles resultados del exp erimento son cono cidos y se pueden describir completamente, pero el resultado del mismo no es cono cido; mientras que en las situaciones ba jo imprecisión, se cono ce el resultado del exp erimento, p ero no es p osible describirlo con precisión. Por tanto, la incertidumbre se mo delará mediante la Teoría de la Probabilidad, mientras que la imprecisión será mo delada mediante la Teoría de los Conjuntos Intuicionísticos. Además, cuando ambas faltas de información aparezcan simultáneamente, se utilizará la Teorí a de las Probabilidades Imprecisas.

Cuando las alternativas a comparar estén definidas ba jo incertidumbre, éstas se mo delarán mediante variables aleatorias. Por tanto, para compararlas se rá necesario utilizar un orden esto cástico. En esta memoria se consideran dos órdenes: la dominancia esto cástica y la preferencia estadística. El primero de ellos es uno de los méto dos más utilizados en la literatura, mientras que el segundo es el méto do óptimo de comparación de variables cualitativas. Para estos méto dos se han estudiado varias propiedades. En particular, si bien es cono cido que la dominancia esto cástica está relacionada con la comparación de las esp eranzas de determinadas trasformaciones de las variables, se prueba que la preferencia estadística está más ligada a otro parámetro de lo calización, la mediana. Además, se han encontrado situaciones ba jo las cuales la dominancia estocástica está relacionada con la preferencia estadística. Estos dos órdenes esto cásticos han sido definidos para comp arar variables aleatorias p or pares. Por esta razón se ha definido una extensión de la preferencia estadística para la comparación simultánea de más de dos variables y se han e studiado varias prop iedades.

Cuando las alte rnativas están defin idas en un marco de incertidumbre e imprecisión, cada una de ellas se mo delará mediante un conjunto de variables aleatorias. Dado que los órdenes esto cásticos comparan variables aleatorias, es necesario realizar su extensión para la comparación de conjuntos de variables. Cuando el orden esto cástico utilizado es la dominancia esto cástica o la preferencia estadística, la comparación de los conjuntos de

variables aleatorias está claramente relacionada con la comparación de elementos propios de la teoría de las probabilidades imprecisas, como pueden ser las p-boxes. Gracias al modelo general que desarrollaremos, se podrán estudiar en particular dos situaciones habituales en los problemas de la teoría de la decisión: la comparación de variables aleatorias bajo utilidades o bajo creencias imprecisas. El primer problema se modelará mediante conjuntos aleatorios, y por lo tanto su comparación se realizará a través de sus conjuntos de selecciones medibles. El segundo problema será modelado mediante un conjunto de probabilidades. Cuando las distribuciones marginales de las variables están definidas bajo imprecisión, la distribución conjunta no se puede obtener mediante el Teorema de Sklar. Por ello, resulta necesario investigar una versión imprecisa de este resultado, que tendrá importantes aplicaciones en los órdenes estocásticos bivariantes definidos bajo imprecisión.

Si las alternativas se definen bajo imprecisión, pero no bajo incertidumbre, éstas se modelarán mediante conjuntos intuicionísticos. Para su comparación se introduce una teoría matemática de comparación de este tipo de conjuntos, dando especial relevancia al concepto de IF-divergencia. Estas medidas de comparación de conjuntos intuicionísticos poseen numerosas aplicaciones, como pueden ser en el reconocimiento de patrones o la teoría de la decisión. Los conjuntos intuicionísticos permiten grados de pertenencia y de no pertenencia, y por ello resultan un buen modelo bipolar. Dado que las probabilidades imprecisas también son utilizadas en el contexto de la información bipolar, se estudiarán las conexiones entre ambas teorías. Estos resultados mostrarán tener interesantes aplicaciones, y en particular permitirán extender la dominancia estocástica para la comparación de más de dos p-boxes.

Abstract

In real life situations it is common to deal with the comparison of alternatives. The alternatives to b e compare d are sometimes defined under some lack of information. Two lacks of information are considered: uncertainty and imprecis ion. Un certainty refers to situations in which the p os sible results of the exp eriment are precisely describ ed, but the exact result of the exp eriment is unknown; imprecision refers to situations in which the result of the exp eriment is known but it cannot be precisely describ ed. In this work, uncertainty is mo delled by means of Probability Theory, imprecision is mo delled by means of IF-set Theory, and th e Theory of Imprecise Probabilities is used when b oth lacks of information hold together.

Alternatives under uncertainty are mo delled by means of random variables. Thus, a sto chastic order is needed for their comparison. In this work two particular sto chastic orders are considered: sto chastic dominance and statistical preference. The former is one of the most usual metho ds used in the literature and the latter is the most adequate metho d for comparing qualitative variables. Some prop erties about such metho ds are investigated. In particular, although sto chastic dominance is related to the exp ectation of some transformation of the random variables, statistical preference is re lated to a different lo cation parameter: the median. In addition, some conditions, related to the copul a that links the random variables, under which sto chastic dominance and statistical preference are related are given. Both sto chastic orders are defined for the pairwise comparison of random variables. Thus, an extension of statistical preference for the comparison of more than two random variables is defined, and its main prop erties are studied.

When the alternatives are defined under uncertainty and imprec ision, each one is represented by a set of random variables. For comparing them, sto chastic orders are extended for the comparison of sets of random variables instead of single ones. When the sto chastic order is either stochastic dominance or statistical preference, the comparison of sets of random variables can b e related to the comparison of elements of th e imprecise probability theory, like p-b oxes. Two particular instances of comparison of sets of random variables, common in de cision making problems, are studied: the comparison of random variables with imprecision on the utilitie s or in the b eliefs. The former situation is mo delled by random sets, and then their sets of measu rable selections are compared, and the second is modelled by a set of probabilities. When there is imprecision in the marginal distributions of the random variables, the joint distribution cannot be obtained from Sklar's Theorem. For this reason, an imprecise version of Sklar's Theorem is given, and its applications to bivariate stochastic orders under imprecision are showed.

Alternatives defined under imprecision, but not under uncertainty, are modeled by means of IF-sets. For their comparison a mathematical theory of comparison of IF-sets is given, focusing on a particular type of measure called IF-divergences. This measure has several applications, like for instance in pattern recognition or decision making. IF-sets are used to model bipolar information because they allow membership and non-membership degrees. Since imprecise probabilities also allow to model bipolarity, a connection between both theories is established. As an application of this connection, an extension of stochastic dominance for the comparison of more than two p-boxes is showed.

1 Introduction

The mathematical modeling of real life experiments can be rendered difficult by the presence of two types of lack of information: uncertainty and imprecision. We speak about uncertainty when the variables involved in the experiment are precisely described but we cannot predict beforehand the outcome of the experiment. This lack of information is usually modelled by means of Probability Theory. On the other hand, imprecision refers to situations in which the result of the experiment is known but it cannot be precisely described. One possible model for this situation is given by Fuzzy Set Theory or any of its extensions, such as the Theory of Intuitionistic Fuzzy Sets or the Theory of Interval-Valued Fuzzy Sets. Of course, there are also situations in which both uncertainty and imprecision appear together. In such cases, we can either combine probability theory and fuzzy sets, or consider the Theory of Imprecise Probabilities.

Fuzzy sets were introduced by Zadeh ([214]) as a more flexible model than crisp sets, which is particularly useful when dealing with linguistic information. A fuzzy set assigns a value to each element on the universe, called membership degree, which is interpreted as the degree in which the element fulfills the characteristic described by the set. Of course, crisp sets are particular cases of fuzzy sets, since every element either belongs (i.e., has membership degree 1) or does not (membership degree equals 0) to the set. Since their introduction, fuzzy sets have become a very popular research topic, and nowadays several international journals, conferences and societies are devoted to them. For a complete study on fuzzy sets, we remit the reader to some usual references like ([71, 101]).

In 1983, Atannasov ([4]) proposed a generalization of fuzzy sets, called the theory of Intuitionistic Fuzzy Sets (IF-sets, for short). In the subsequent years he continued developing his idea ([5, 7]), and now it has become a commonly accepted generalization of fuzzy sets. While fuzzy sets give a degree of membership of every element to the set, an IF-set assigns both a degree of membership and a degree of non-membership of any element to the set, with the natural restriction of that their sum must not exceed 1. Every IF-set has a degree of indeterminacy or uncertainty, that is, one minus the sum of the degrees of membership and non-membership. In this sense we can see that every fuzzy set is in particular an IF-set, since the non-membership degree of the fuzzy set is one minus its membership degree: the indeterminacy degree of a fuzzy set equals zero. For this reason IF-sets have become a very useful tool in order to model situations in which human answers are present: *yes*, *no* or *does not apply*, like for example human votes ([8]). On the other hand, Zadeh also proposed several generalizations of fuzzy sets ([216]). In particular, he introduced interval-valued fuzzy sets (IVF-sets, for short): when the membership degree of an element to the set cannot be precisely determined, it assigns an interval that contains the real membership degree. Although IF-sets and IVF-sets differ on the interpretation, they are formally equivalent (see [30]). These theories have been applied to different areas, like decision making ([194]), logic programming ([9, 10]), medical diagnosis ([48]), pattern recognition ([92]) and interesting theoretical developments are still being made (see for example [68, 97, 120]).

The second pillar of this dissertation is the theory of Imprecise Probabilities. Imprecise Probability is a generic term that refers to all mathematical models that serve as an alternative and a generalization to probability models in cases of imprecise knowledge. It includes possibility measures ([217]), Choquet capacities ([39]), belief functions ([187]) or coherent lower previsions ([205]), among others. One model that will be of particular interest for us is that of p-boxes. A p-box ([75]) is determined by an ordered pair of functions called lower and upper distribution functions, and it is given by all the distribution functions bounded between them. Troffaes et al. ([198, 201]) have investigated the connection between p-boxes and coherent lower probabilities ([205]). In particular, they found conditions under which a p-box defines a coherent lower probability. In some recent papers ([64, 65, 199, 200]) the authors have explored the connection between p-boxes and other usual models included in the theory of imprecise probabilities, such as possibilities, belief functions or clouds ([168]), among others.

This memory deals with the comparison of alternatives under lack of information. As we mentioned before, we shall consider the comparison under uncertainty, imprecision or both. On the one hand, alternatives under uncertainty are modelled by means of random variables. Random variables are one tool of the probability theory that provide a formal background to model non-deterministic situations, that is, situations where randomness is present. The comparison of random variables is a long standing problem that has been tackled from many points of view (see among others [18, 90, 98, 106, 188, 192, 210]). Its practical interest is clear since many real life processes are modelled by random variables. The procedures of comparison are referred to as stochastic orders. Indeed, stochastic ordering is a very popular topic within Economics ([11, 109]), Finance ([110, 173]), Social Welfare ([77]), Agriculture ([95]), Soft Computing ([180, 183]) or Operational Research ([171]), among others.

One classical way of pairwise ordering random variables is stochastic dominance ([108, 208]), a generalization of the expected utility model. First degree stochastic dominance, that seems to be the most widely used method, orders random variables by comparing their cumulative distribution functions (or their survival functions). Its main drawback is that it imposes a very strong condition to get an order, so many pairs of random variables are deemed incomparable. Because of this fact, a second definition, called second degree stochastic dominance is also used, specially in Economics ([98, 139]). Although less restrictive, it still does not establish a complete order between random variables. In fact, we can weaken progressively the notion of stochastic dominance, and talk of stochastic dominance of n-th order.

One interesting alternative stochastic order is statistical preference, particularly when comparing qualitative random variables, taking into account the results by Dubois et al. ([67]). Although it was introduced by De Schuymer et al. ([55, 57]), it is possible to find similar methods in the literature (see [25, 26, 210]). The notion of statistical preference is based on a probabilistic relation, also called reciprocal relation ([21]), that measures the degree of preference of one random variable over the other one. Furthermore, since statistical preference depends on the joint distribution of the random variables, it depends on the copula ([166]) that links them. Recall that from Sklar's Theorem ([189]) it is known that for any two random variables there exists a function, called copula, that allows to express the joint cumulative distribution function in terms of the marginals. Then, statistical preference depends on such copula. The main drawback of this method is its lack of transitivity. Some authors have been investigating which kind of transitivity properties are satisfied by statistical preference, and in particular they focused on cycle-transitivity (see [14, 15, 16, 49, 54, 56, 58, 121, 122]).

When the alternatives to be compared are defined under both uncertainty and imprecision, the problem of comparing sets of random variables arises. Here we understand the set of random variables from an epistemic point of view: we assume that the set of random variables contains the true random variable, but such random variable is unknown ([73]). This situation is not uncommon in decision making under uncertainty, where there is vague or conflicting information about the probabilities or the utilities associated to the different alternatives. We may think for instance of conflicts among the opinions of several experts, limits or errors in the observational process, or simply partial or total ignorance about the process underlying the alternatives. In any of such cases, the elicitation of an unique probability/utility model for each of the alternatives may be difficult and its use, questionable.

Indeed, one of the solutions that have been proposed for situations like this is to consider a robust approach, by means of a set of probabilities and utilities. The use of this approach to compare two alternatives is formally equivalent to the comparison of two sets of alternatives, those associated to each possible probability-utility pair. Hence, it becomes useful to consider comparison methods that allow us to deal with sets of alternatives instead of single ones.

However, the way to compare of sets of alternatives is no longer immediate: we may compare all possibilities within each of the sets, or also select some particular elements of each set, to take into account phenomena of risk aversion, for instance. This gives rise to a number of possibilities. Moreover, even in the simpler case where we choose one alternative from each set, we must still decide which criterion we shall consider to determine the preferred one. There is quite an extensive literature on how to deal with imprecise beliefs and utilities when our choice is made by means of an expected utility model ([12, 165, 178, 186]). However, the problem has almost remained unexplored for other choice functions. For this reason, we shall extend stochastic orders for the comparison of sets of random variables, and we shall see that the proposed extension is connected to the imprecise probability theory.

The last situation to be studied is the comparison of alternatives under imprecision but without uncertainty. In this case the alternatives will be described by means of IF-sets. Within fuzzy set theory, several types of measures of comparison have been defined, with the goal of quantifying how different two fuzzy sets are. The more usual measures of comparison are dissimilarities ([119]), dissimilitudes ([44]) and divergences ([159]). Other authors, like Bouchon-Meunier et al. ([27]), defined a general axiomatic framework for the comparison of fuzzy sets, that include the aforementioned measures as particular cases. Montes ([159]) made a complete study of the divergences as a measure of comparison of fuzzy sets. In particular, she introduced a particular kind of divergences, called local divergences, that have proven to be very useful.

Distances between fuzzy sets are also important for many practical applications. For instance, Bhandari et al. ([22]) proposed a divergence measure for fuzzy sets inspired by the notion of divergence between two probability distributions, and used this fuzzy divergence measure in the framework of image segmentation. Several other attempts within the same field have been considered ([23, 34, 74]). For instance, the fuzzy divergence measure of Fan and Xie is based (unlike the proposal of Bhandari and Pal) on the exponential entropy of Pal and Pal ([175]); the same spirit is followed in [34].

However, in the framework of IF-sets only the notion of distance as well as several examples of IF-dissimilarities have been given (see for example [36, 37, 85, 89, 92, 111, 113, 114, 138, 193]). Nevertheless, the need for a formal mathematical theory of comparison of IF-sets still persists.

Furthermore, IF-sets are a very useful tool to represent bipolar information: the membership and non-membership degree of every element to the set. Since bipolar models are also being studied within the framework of imprecise probabilities (see for instance [64, 65, 72, 73]), it becomes natural to investigate the connection between both approaches to the modeling of bipolar information.

The rest of the work is organized as follows. Chapter 2 introduces the basic notions that will be necessary along the work. In the first part we deal with stochastic orders, focusing on stochastic dominance, that is based on the comparison of the cumulative distribution functions of the random variables, and statistical preference, that is based on a probabilistic relation and makes use of the joint distribution. In order to express this joint distribution as a function of the marginals, we need to introduce some notions of the theory of copulas. Then, we make a brief introduction to the theory of imprecise probabilities. On the first part we define coherent lower previsions and we recall the basic results we shall use later on. Then, we focus on particular cases of coherent lower probabilities: n-monotone capacities, belief functions, possibility measures and clouds. We also define random sets and show their connections with imprecise probability theory. Finally, we make an overview of IF-sets theory. First, we explain the semantic differences between IF-sets and IVF-sets and show that both theories are formally equivalent. Then, we introduce the basic operations between these sets.

In Chapter 3 we investigate the comparison of alternatives under uncertainty, that will be modelled by means of random variables. Although some stochastic orders like stochastic dominance have already been widely explored in the literature, this is not the case for statistical preference. For this reason, we devote Section 3.1 to investigate the main properties of this relation, and we compare them to the ones of stochastic dominance ([149, 154]). While stochastic dominance has a well-known characterization in terms of the comparison of the expectations of adequate transformations of the random variables, there is not a characterization of statistical preference. For this aim, we investigate a possible characterization in terms of expectations ([150, 153]) and in terms of a different location parameter: the median ([148, 163]).

Although statistical preference and stochastic dominance are not related in general, in Section 3.2 we look for conditions under which first degree stochastic dominance implies statistical preference ([150]). Obviously, since statistical preference depends on the copula that links the variables, these conditions are related to such copula. Furthermore, we find that in some of the usual probability distributions, like Bernoulli, uniform, normal, etc, both stochastic dominance and statistical preference are equivalent for independent random variables ([151]).

We have already mentioned the lack of transitivity of statistical preference, which renders it unsuitable for comparing more than two random variables. In order to overcome this problem, we introduce in Section 3.3 an extension of statistical preference that preserves its philosophy and allows the comparison of more than two random variables ([140, 142]). We explore this new notion and give several properties that relate it to the classical notion of statistical preference. In order to illustrate the applicability of our results, Section 3.4.1 puts forward two different applications. We first use both stochastic dominance and statistical preference to compare fitness values associated to the output of genetic fuzzy systems ([143, 152, 162]), and then we use the generalization of statistical preference on a decision-making problem with linguistic variables.

In Chapter 4 we consider the comparison of alternatives under both uncertainty and imprecision. As we have already mentioned, in that case we model the alternatives by means of sets of random variables instead of single ones. We start in Section 4.1 by extending binary relations that are used to the comparison of random variables to the comparison of sets of random variables. This gives rise to six possible ways of comparing sets of random variables. In particular, we focus on the case where such binary relation is either stochastic dominance or statistical preference. We shall see that the use of stochastic dominance as binary relation is clearly connected to the comparison of the p-boxes associated with the sets of random variables ([134, 155, 157]). We shall consider two particular cases in Section 4.2: the comparison of two random variables with imprecise utilities and the comparison of two random variables with imprecise beliefs ([156]). The former is modelled by means of random sets, and their comparison is made by means of their associated sets of measurable selections. In the latter, the imprecise beliefs are modelled by means of a set of probabilities in the initial space, instead of a single one. In this situation we can also define a set of random variables for each alternative. Then, both situations are particular cases of the more general situation studied in Section 4.1.

When there is imprecision about the probability of the initial space, the joint distribution of the random variables is also imprecisely determined. Because of this, it seems reasonable to investigate how the bivariate distribution, and in particular the bivariate cumulative distribution function, can be determined. We shall investigate the properties of bivariate p-boxes and how they can define a coherent lower probability ([135]). One particular instance where the joint distribution naturally arises is when dealing with copulas. Recall that copulas allow to determine the joint distribution function in terms of the marginals. However, when the marginal distribution functions are imprecisely described by means of p-boxes, it is unclear how to determine the joint distribution, and bivariate p-boxes prove to be a useful tool. In particular we show that, by considering an imprecise version of copulas it is possible to extend Sklar's Theorem to an imprecise framework ([176]).

Section 4.4 shows several applications of the results from Chapter 4. One possible application is the comparison of Lorenz Curves ([3, 11]), that represent the inequalities within countries/regions. Using our results, it is possible to compare sets of regions by means of stochastic dominance. Furthermore, imprecise stochastic dominance also allows to compare survival rates of different cancer grouped by sites. We conclude the chapter showing another application in decision making.

In Chapter 5 we investigate how to compare alternatives under imprecision. The alternatives are modelled by means of IF-sets, and we propose methods for comparing IF-sets. In Section 5.1 we recall the comparison measures that can be found in the literature: IF-dissimilarities and distances for IF-sets. We also introduce IF-divergences and IF-dissimilitudes ([141]). We investigate the relationships among these measures and we justify that our preference for IF-divergences in that they impose stronger conditions, avoiding thus counterintuitive examples ([145, 161]). We also try to define a general measure of comparison of IF-sets as done by Bouchon-Meunier et al. ([27]) for fuzzy sets. This allows us to define a general function that contains IF-dissimilarities, IF-divergences and distances as particular cases ([158]). Then we introduce a particular type of IF-divergences, that are those that satisfy a local property. We investigate their properties and give several examples ([147]). We conclude the section studying the connection between IF-divergences and divergences for fuzzy sets. In particular, we show how we can define IF-divergences from divergences for fuzzy sets and, conversely, how to build divergences for fuzzy sets from IF-divergences ([146]).

Since both imprecise probabilities and IF-sets are used to mo del bip olarity, we investigate in Section 5.2 the connection b etween both approaches. Weestablish that when IF-sets are defined in a probability space, they can be interpreted as random sets, and this allows to connect them with impreci se probabilities, since it is p ossible to define acredal set and a lower and upp er probability. Weinvestigateunderwhich conditions the probabilistic information enc o ded by the credal set is the same than the one of the set of measurable selections.We also inve stigate the relationship b etween our approach and otherworksin the literature, like the one of Grzegorzewski and Mrowka ([86]).

Weconclude the chapter showing several applications of the results. Onthe one hand we show how IF-divergences can be applied to decision making and pattern recognition. On the other hand, we explain how the connection b etwe en IF-sets and imprecise probabilities allows us to prop ose a generalization of sto chastic dominance to the comparison of more than two p-b oxes, and we illustrate our metho d comparing at the same time sets of Lorenz Curves.

We conclude this dissertation with some final remarks and adiscussion of the most imp ortant future lines of research.

2 Basic concepts

In this chapter, we intro duce the main notions that shall be employed in the rest of the work. We start by p roviding th e definition of binary relations as comparison metho ds for random variables. Later, weconsidertheparticularcaseswherethebinaryrelationis either sto chastic dominance or statistical preference, which are the two main sto chastic orders we shall consider here.

Afterwards we make a brief intro duction to Imprecise Probability theory, that shall b e useful when we want to compare sets of random variables.Toconclude thechapter, we recall the notion of intuitionistic fuzzy sets, that we shall use mo del situations where sets cannot be precisely describ ed.

2.1 Stochastic orders

Stochastic orders are metho ds that determine a (total or partial) order on any given set of random quantities. Although several methods have b een prop osed in the last years (see for instance [139, 188]), here we shall fo cus on two particular cases: sto chastic dominance and statistical preference. The former is p ossibly the most widespread metho d in the literature, andthe latter isparticularly useful whencomparing qualitative variables, taking into account the axiomatization established by Dub ois et al. ([67]).

Throughout, randomvariablesare denotedby X, Y, Z, . . ., or X_1,X_2, . . ., and their asso ciated cumulative distribution functions are denoted F_X, F_Y, F_Z, . . ., or F_{X_1},F_{X_2}, . . ., resp ectively. We shall also assume that the random variables to be compared are definedon thesameprobability space.

Given two random variables X and Y definedfromthe probability space $(\Omega, \mathcal{A}, P)$ to an ordered space $(\Omega, \mathcal{A})$ (which in most situations will be the set of real numb ers), abinary relation is usedto comparethe variables. Then, X Y means that X is at leastas preferable as Y. This corresp onds to a weak preference relation; from it a strict preference relation, indifference and also incomparable relation can also be defined:

Definition 2.1 *Considertwo randomvariables X and Y and a binary relation used to compare them.*

- *X is* strictlypreferred *to Y with respectto , and isdenotedby X Y, if X Y but Y X.*
- *X and Y* areindifferent *with respect to , and itis denotedby $X \equiv Y$, if X Y and Y X.*
- *X and Y* areincomparable*with respect to , andit is denotedby X Y, if X Y and Y X.*

Then, if D denotes a setof random variables, accordingto [179], $(D, \ , \equiv, \ )$ formsa preference structure. Inparticular, iftherelation is complete*, that is, if there is not incomparabilitybetween the random variables, then $(D, \ , \equiv)$ forms apreference structure without incomparable elements.*

One instance of binary relation is the comparis on of the exp ectations of the random variables, so that X Y if and only if $E(X) \geq E(Y)$. This is also anexampleof a non-complete relation, because the comparison cannot be made when the exp ectation of the variable do es not exist.

In the remainder of this section we intro duce the definitions and notations that we shall use in the following chapters. Sp ecifically,we consi der the case in which the binary relation is either sto chastic dominance or statistical preference. With res p ect to the first one, we recall the main typ es of sto chastic dominance and some ofits most imp ortant prop erties, suchasits characterizationbymeansof the comparison of the adequate exp ectations.Then, weprovideanoverviewonstatisticalpreference: we rec all its definition and we also discuss briefly its main advantages as a sto chastic order.

2.1.1 Stochastic dominance

Sto chastic dominance is one ofthe most used metho ds for the pairwise comparison of randomvariables we can find in the literature. Besides to the usual economic interpretation (see [110]), this notion has also b ee n applied in other frameworks such as Finance ([109]), So cialWelfare ([11]), Agriculture ([95]) or Op erations Research ([171]), among others. We next recall its definition and basic notions related to th em, and also its main prop erties.

Stochastic dominance is a metho d based on the comparison of the cumulative distribution functions of therandomvariables.

Definition 2.2 *Let X and Y betworeal-valuedrandom variables, and let F_X and F_Y denote their respective cumu lative distribution functions. X* sto chastically dominates *Y*

by the first degree, *or simply stochastically dominates, when no confusion is possible, and it is denoted by* $X \ _{FSD} Y$*, if it holds that*

$$F_X(t) = P(X \leq t) \leq P(Y \leq t) = F_Y(t) \text{ for every } t \quad \mathbb{R}. \tag{2.1}$$

One of the most important drawbacks of this definition is that (first degree) stochastic dominance is a non-complete relation, that is, it is possible to find random variables X and Y such that neither $X \ _{FSD} Y$ nor $Y \ _{FSD} X$, as we can see in the following example.

Example 2.3 *Consider two random variables X and Y such that X follows a Bernoulli distribution with parameter 0.6 and Y takes a fixed value c $(0, 0.6)$ with probability 1. Then, there is not first degree stochastic dominance between them:*

$$F_X(0) = 0.4 > 0 = F_Y(0) \text{ but } F_X(c) = 0.4 < 1 = F_Y(c).$$

According to Definition 2.1, from this preference relation we can also define the strict stochastic dominance, the indifference and, as we have just seen, the incomparability relations:

- X stochastically dominates Y *strictly*, and denote it by $X \ _{FSD} Y$, if and only if $F_X \leq F_Y$ and there is some t $[0, 1]$ such that $F_X(t) < F_Y(t)$.
- X and Y are *stochastically indifferent*, and denote it by $F_X \equiv_{FSD} F_Y$, if and only if they have the same distribution (usually denoted by $X \overset{d}{=} Y$).
- X and Y are *stochastically incomparable*, and denote it by $X \quad Y$, if there are t_1 and t_2 such that $F_X(t_1) > F_Y(t_2)$ and $F_Y(t_2) > F_X(t_2)$.

Remark 2.4 *Here we have chosen the notation $_{FSD}$ because it is the most frequent in the literature. However, (first degree) stochastic dominance has also been denoted by $_1$, as in [55], or by $\geq_{st}$, as in [188]. In that case, the authors used the name* stochastic order *instead of first degree stochastic dominance.*

As we see from its definition, (first degree) stochastic dominance only focuses on the marginal cumulative distribution functions, and its interpretation is the following: if $X \ _{FSD} Y$, then $F_X(t) \leq F_Y(t)$ for any t, or equivalently, $P(X > t) \geq P(Y > t)$ for any t. That is, we impose that at every point the probability of X to be greater than such point is greater than the probability of Y to be greater than the same point. Thus, X assigns greater probability to greater values. Figure 2.1 shows its graphical interpretation. Here, we can see how F_X is always below or at the same level than F_Y.

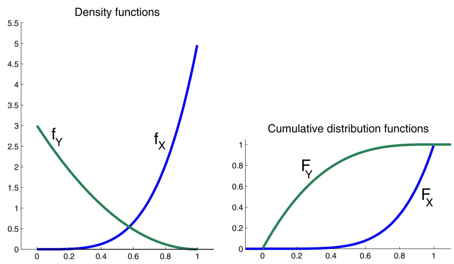


Figure 2.1: Example of first degree stochastic dominance: $X \succeq_{FSD} Y$

From an economic point of view, the interpretation is that the decision between the two random variables is rational, in the sense that for any threshold of profit the probability of going above this threshold is greater with the preferred variable ([110]).

The main drawback of this definition is that the inequality in Equation (2.1) is quite restrictive. There are many pairs of cumulative distribution functions that do not satisfy this inequality in any sense and therefore, the associated random variables cannot be ordered. This is the reason why we can consider other (weaker) degrees of stochastic dominance. Let us now introduce the second degree stochastic dominance.

Definition 2.5 *Let X and Y be two real-valued random variables whose cumulative distribution functions are given by F_X and F_Y, respectively. X* sto chastically dominates Y by the second degree*, and it is denoted by $X \succeq_{SSD} Y$, if it holds that:*

$$\int_{-\infty}^{t} F_X(x)\,\mathrm{d}(x) \le \int_{-\infty}^{t} F_Y(y)\,\mathrm{d}(y) \text{ for every } t \in \mathbb{R}. \tag{2.2}$$

As in Definition 2.2, we can also introduce the strict second degree stochastic dominance ($\succ_{SSD}$), the indifference ($\equiv_{SSD}$) and the incomparable ($\parallel_{SSD}$) relations.

Note that, similar to Example 2.3, we can also see that incomparability is possible when dealing with second degree stochastic dominance.

Example 2.6 *Consider the same random variables of Example 2.3. For these variables, the functions G_X^2 and G_Y^2 are defined by:*

$$G_X^2(t) = \begin{cases} 0 & \text{if } t < 0. \\ 0.4t & \text{if } t \in [0,1). \\ t - 0.6 & \text{if } t \ge 1. \end{cases} \qquad G_Y^2(t) = \begin{cases} 0 & \text{if } t < c. \\ t - c & \text{if } t \ge c. \end{cases}$$

Then, X and Y are not ordered by means of the second degree stochastic dominance since:

$$G_X^2\left(\frac{c}{2}\right) = 0.2c > 0 = G_Y^2\left(\frac{c}{2}\right) \text{ but } G_X^2(1) = 0.4 < 1 - c = G_Y^2(1),$$

since $c < 0.6$.

Remark 2.7 *Other authors (see for example [188]) call this method concave order, and they denote it by* $\geq_{cv}$. *It is also sometimes denoted by* $\succeq_2$ *([55]).*

As we can see in Figure 2.2, when $X \succeq_{SSD} Y$, for any fixed t, the area below F_X until t is lower than the are below F_Y until t. This means that the X gathers more accumulated probability at greater points than Y.

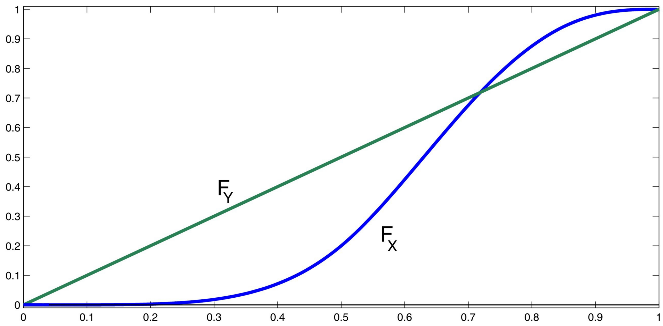


Figure 2.2: Example of second degree stochastic dominance: $X \succeq_{SSD} Y$.

From an economic point of view, second degree stochastic dominance means that the decision maker prefers the alternative that provides a bigger profit but also with less risk. That is, it is a rationality criterion under risk aversion (see [110]).

Similarly to Definitions 2.2 and 2.5, stochastic dominance can be defined for every degree n by relaxing the conditions in Equations (2.1) and (2.2).

Definition 2.8 *Let* X *and* Y *be two real-valued random variables with cumulative distribution functions* F_X *and* F_Y*, respectively.* X stochastically dominates Y by the n-th degree, *for* $n \geq 2$*, and it is denoted by* $X \succeq_{nSD} Y$*, if it holds that:*

$$G_X^n(t) = \int_{-\infty}^{t} G_X^{n-1}(x)d(x) \leq \int_{-\infty}^{t} G_Y^{n-1}(y)d(y) = G_Y^n(t) \quad \forall t \in \mathbb{R}, \tag{2.3}$$

where $G_X^1 = F_X$ *and* $G_Y^1 = F_Y$*. In particular, this definition becomes the second degree stochastic dominance when* $n = 2$.

Again, following the notation of Definition 2.1, we can introduce the strict n-th degree stochastic dominance ($\succ_{nSD}$), the indifference ($\equiv_{nSD}$) and the incomparability ($\bowtie_{nSD}$)

relations. Then, if D denotes a set of random variables, $(D, {}_{nSD}, \equiv_{nSD}, {}_{nSD})$ forms a preference structure for any $n \geq 1$.

Clearly, first degree stochastic dominance imposes a stronger condition than second degree stochastic dominance, as we can see from Equations (2.1) and (2.2). Moreover, if we compare Equations (2.1) and (2.3) we deduce that first degree stochastic dominance is stronger than the n-th degree stochastic dominance for every n. Indeed, it is known that the n-th degree stochastic dominance is stronger than the m-th degree stochastic dominance for any $n<m$:

$$X \quad _{nSD} \; Y \qquad X \geq_{mSD} Y \text{ for every } n < m, \tag{2.4}$$

while the converse does not hold in general.

Remark 2.9 *Stochastic dominance is a reflexive and transitive relation. However, since two different random variables may induce the same distribution, it is not antisymmetric. Moreover, as we have already noted, it is not complete because it allows incomparability.*

One of the most important properties of stochastic dominance is its characterization by means of the expectation. Specifically, each of the types of stochastic dominance we have introduced can be characterized by the comparison of the expectations of adequate transformations of the variables considered.

Theorem 2.10 ([109, 139]) *Let X and Y be two random variables. For first and second degree stochastic dominance it holds that:*

- $X \quad _{FSD} \; Y$ *if and only if* $E[u(X)] \geq E[u(Y)]$ *for every increasing function* $u: \mathbb{R} \to \mathbb{R}$.
- $X \quad _{SSD} \; Y$ *if and only if* $E[u(X)] \geq E[u(Y)]$ *for every increasing and concave function* $u: \mathbb{R} \to \mathbb{R}$.

A function $u: \mathbb{R} \to \mathbb{R}$ is called n-monotone ([39]) if it is n-differentiable and for any $m \leq n$ and it fulfills $(-1)^{m+1} u^{(m)} \geq 0$. Then, if U_n denotes the set of n-monotone functions, the following general equivalence holds:

$$X \quad _{nSD} \; Y \qquad E[u(X)] \geq E[u(Y)] \text{ for every } u \quad U_n. \tag{2.5}$$

In fact, from the proof of Theorem 2.10, it can be derived that:

$$X \quad _{nSD} \; Y \qquad E[u(X)] \geq E[u(Y)] \text{ for any } u \quad U_n, \tag{2.6}$$

where U_n denote the set of n-monotone and bounded functions $u: \mathbb{R} \to \mathbb{R}$.

Equation (2.4) can also be derived from this result, since every n-monotone function is also m-monotone for any $m \leq n$.

Remark 2.11 *The characterization of the second degree stochastic dominance, based on the comparison of the mean of the concave and increasing functions, explains the nomenclature* concave order *mentioned in Remark 2.7.*

To conclude this paragraph, we list some interesting properties of first degree stochastic dominance that shall be useful in the next chapter.

Proposition 2.12 ([139, Theorem 1.2.13]) *If X and Y are real-valued random variables such that $X \;{}_{FSD}\; Y$ and $\phi: \mathbb{R} \to \mathbb{R}$ is a increasing function, then $\phi(X) \;{}_{FSD}\; \phi(Y)$.*

Proposition 2.13 ([139, Theorem 1.2.17]) *Let $\{X_i, Y_i : i = 1, \ldots, n\}$ be independent and real-valued random variables. If $X_i \;{}_{FSD}\; Y_i$ for $i = 1, \ldots, n$, then $X_1 + \ldots + X_n \;{}_{FSD}\; Y_1 + \ldots + Y_n$.*

Proposition 2.14 ([139, Theorem 1.2.14]) *Given the random variables $X, X_1, X_2, \ldots, Y, Y_1, Y_2, \ldots$ such that $X_n \xrightarrow{L} X$ and $Y_n \xrightarrow{L} Y$, if $X_n \;{}_{FSD}\; Y_n$ for every n, where $\xrightarrow{L}$ denotes the convergence in distribution, then $X \;{}_{FSD}\; Y$.*

As a consequence of the previous result, first degree stochastic dominance is preserved by four kinds of converge: distribution, probability, m^{th}-mean and almost sure.

For a more complete study on stochastic orders, we refer to [62, 109, 139, 188, 192].

2.1.2 Statistical preference

In the previous subsection we have mentioned that stochastic dominance is a pairwise comparison method that has been used in several areas, always with successful results. However, this method also presents some drawbacks: on the one hand, it is a non-complete crisp relation. This means that it is possible to find pairs of random variables such that n-th degree stochastic dominance does not order them for any n. Furthermore, stochastic dominance does not allow to establish degrees of preference. In fact, there are only three possibilities: either one random variable is preferred to the other, or they are indifferent or incomparable. in addition, it is a method with a high computational cost, since the n-th degree stochastic dominance requires the computation of $2(n-1)$ integrals.

These drawbacks made De Schuymer et al. ([55, 57]) introduce a new method for the pairwise comparison of the random variables, based on a probabilistic relation.

Definition 2.15 ([21]) *Given a set of alternatives D, a* probabilistic or reciprocal relation *Q is a map $Q: D \times D \to [0,1]$ such that $Q(a,b) + Q(b,a) = 1$ for any alternatives $a, b \;\; D$.*

In our framework, the set of alternatives D is considered to be made by random variables defined on the same probability space $(\Omega, \mathcal{P}(\Omega), P)$ to an ordered space $(\Omega, \mathcal{A})$. The probabilistic relation over D is defined (see [55, Equation 3]) by:

$$Q(X, Y) = P(X > Y) + \frac{1}{2} P(X = Y), \tag{2.7}$$

where (X, Y) $D \times D$ and P denotes the joint probability of the bidimensional random vector (X, Y). Clearly, Q is a probabilistic relation: it takes values in $[0, 1]$ and $Q(X, Y) + Q(Y, X) = 1$:

$$Q(X, Y) + Q(Y, X) = P(X > Y) + \frac{1}{2} P(X = Y) + \frac{1}{2} P(X = Y) + P(Y > X) = 1.$$

The above definition measures the preference degree of a random variable X over another random variable Y, in the sense that the greater the value of $Q(X, Y)$, the stronger the preference of X over Y. Hence, the closer the value $Q(X, Y)$ is to 1, the greater we consider X with respect to Y; the closer $Q(X, Y)$ is to 0, the greater we consider Y to X; and if $Q(X, Y)$ is around 0.5, both alternatives are considered indifferent. This fact can be seen in Figure 2.3.

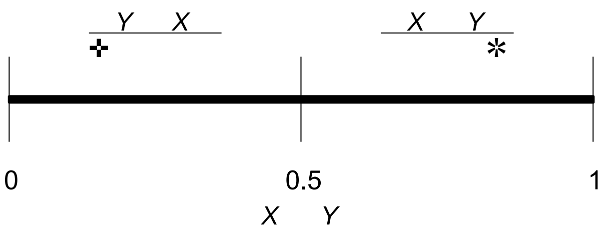


Figure 2.3: Interpretation of the reciprocal relation Q.

Statistical preference is defined from the probabilistic relation Q of Equation (2.7) and it is the formal interpretation of that relation.

Definition 2.16 ([55, 57]) *Let X and Y be two random variables. It is said that:*

- X *is* statistically preferred *to* Y*, and it is denoted by* X $_{SP}$ Y*, if* $Q(X, Y) \geq \frac{1}{2}$.

Also, according to Definition 2.1:

- X *and* Y *are* statistically indifferent*, and it is denoted by* $X \equiv_{SP} Y$*, if* $Q(X, Y) = \frac{1}{2}$.
- X *is* strictly statistically preferred *to* Y*, and we denote it* X $_{SP}$ Y*, if* $Q(X, Y) > \frac{1}{2}$.

Note that statistical preference does not allow incomparability, so $(D, \ _{SP}, \equiv_{SP})$ constitutes a preference structure without incomparable elements.

Remark 2.17 *Statistical preference is a reflexive and complete relation. However, it is neither antisymmetric not transitive, as we shall see in Section 3.3.*

It is possible to give a geometrical interpretation to the concept of statistical preference. As we can see in Figure 2.4, given two continuous and independent random variables, $X \ _{SP} Y$ if and only if the volume enclosed under the joint density function in the half-space $\{(x, y) \mid x > y\}$ is larger than the volume enclosed in the half-space $\{(x, y) \mid x < y\}$.

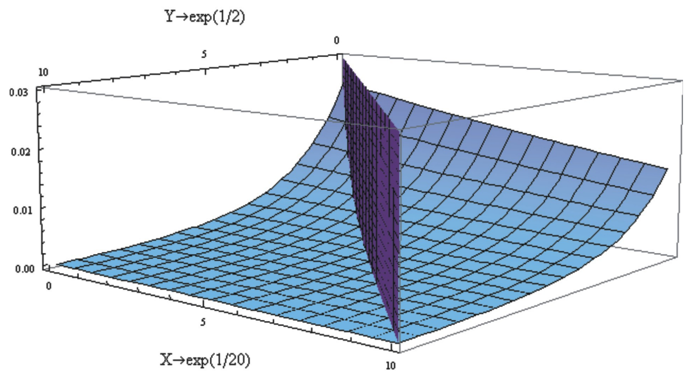


Figure 2.4: Geometrical interpretation of the statistical preference: $X \ _{SP} Y$.

Note that $X \ _{SP} Y$ means that X outperforms Y with a probability at least 0.5. Hence, statistical preference provides an order between the random variables and a preference degree. This is illustrated in the following example.

Example 2.18 *Consider two random variables X, Y such that X follows a Bernoulli distribution $B(p)$ with parameter $p \ (0, 1)$ and Y follows a uniform distribution $U(0, 1)$ in the interval $(0, 1)$. It is immediate that:*

$$Q(X, Y) = P(X > Y) = P(X = 1) = p.$$

Therefore, when $p \geq \frac{1}{2}$, X is statistically preferred to Y with degree of preference p, and the greater the value of p, the most preferred X is to Y.

One imp ortant remark is that statistical preference fordegenerate random variablesis equivalent to the order b etween real numb ers, and in that case the prefere nce degree is always 0, 1or $\frac{1}{2}$.

Remark 2.19 *Considertwo random variables X and Y. The former takes the value c_X with probability 1 and the secondtakes thevalue c_Y with probability 1. Assume that $c_X > c_Y$:*

$$P(X > Y) = P(X = c_X) = 1 \quad Q(X, Y) = 1 \text{ and } X \;_{SP}\; Y.$$

On the other hand, if $c_X = c_Y$, it holds that:

$$P(X = Y) = P(X = c_X, Y = c_Y) = 1 \quad Q(X, Y) = \frac{1}{2} \text{ and } X \equiv_{SP} Y.$$

Then, it holds that:

$$X \;_{SP}\; Y \quad c_X > c_Y \text{ and } X \equiv_{SP} Y \quad c_X = c_Y.$$

A first, but also trivial result ab out statistical preference is the following.

Lemma 2.20 *Given tworandom variables X and Y, it holds that:*

$$\begin{aligned} X \;_{SP}\; Y \quad Q(X, Y) \geq Q(Y, X) \quad & P(X \geq Y) \geq P(Y \geq X) \\ & P(X > Y) \geq P(Y > X). \end{aligned}$$

Pro of By definition, $X \;_{SP}\; Y$ if andonly if $Q(X, Y) \geq \frac{1}{2}$. Since Q is aprobabilistic relation, $Q(X, Y) + Q(Y, X) = 1$. Then:

$$Q(X, Y) \geq \frac{1}{2} \quad Q(X, Y) \geq \frac{1}{2}(Q(X, Y) + Q(Y, X)) \quad Q(X, Y) \geq Q(Y, X).$$

Letus now prove the remaining equivalences.

$$\begin{aligned} X \;_{SP}\; Y \quad & Q(X, Y) \geq Q(Y, X) \\ & P(X > Y) + \tfrac{1}{2}P(X = Y) \geq P(Y > X) + \tfrac{1}{2}P(X = Y) \\ & P(X > Y) \geq P(Y > X). \end{aligned}$$

Moreover:

$$\begin{aligned} X \;_{SP}\; Y \quad & P(X > Y) \geq P(Y > X) \\ & P(X > Y) + P(X = Y) \geq P(Y > X) + P(X = Y) \\ & P(X \geq Y) \geq P(Y \geq X). \end{aligned}$$

Similar equivalences can b e proved for the strict statistical preference:

$$\begin{aligned} X \;_{SP}\; Y \quad Q(X, Y) > Q(Y, X) \quad & P(X \geq Y) > P(Y \geq X) \\ & P(X > Y) > P(Y > X). \end{aligned}$$

Remark 2.21 *One context where statistical preference appears naturally is that of decision making with qualitative random variables. Dubois et al. showed in [67] that given two random variables* $X, Y : \Omega \to \Omega$, *where* $(\Omega, \succeq_\Omega)$ *is an ordered qualitative scale, then, given a number of rationality axioms over our decision rule, the choice between* X *and* Y *must be made by means of* the likely dominance rule, *which says that* X *is preferred to* Y *if and only if* $[X \succ_\Omega Y] \succeq [Y \succ_\Omega X]$, *where:*

$$[X \succ_\Omega Y] = \{\omega \in \Omega : X(\omega) \succ_\Omega Y(\omega)\} \text{ and}$$
$$[Y \succ_\Omega X] = \{\omega \in \Omega : Y(\omega) \succ_\Omega X(\omega)\},$$

where $\succeq$ *is a binary relation on subsets of* Ω. *One of the most interesting cases is that where* $\succeq$ *is determined by a probability measure* P, *so* $A \succeq B \Leftrightarrow P(A) \geq P(B)$. *Then, using Lemma 2.20,* X *is preferred to* Y *if and only if* $X \succeq_{SP} Y$.

We conclude that, according to the axioms considered in [67], statistical preference is the optimal method for comparing qualitative random variables defined on a probability space.

Remark 2.22 *A related notion to statistical preference is that of* probability dominance *considered in [210]:* X *is said to dominate* Y *with probability* $\beta \geq 0.5$, *and it is denoted by* $X\beta Y$, *if* $P(X > Y) \geq \beta$. *This definition has an important drawback with respect to statistical preference, which is that incomparability is possible for every* $\beta \geq 0.5$. *For instance, this is the case of random variables* X *and* Y *satisfying* $P(X = Y) > 0.5$.

In [2], X *is called preferred to* Y *in the precedence order when* $P(X \geq Y) \geq \frac{1}{2}$. *The drawback of this notion is that indifference is possible although* $P(X > Y) > P(Y > X)$, *for instance when* $P(X = Y) \geq \frac{1}{2}$.

From Lemma 2.20 we know that $X \succeq_{SP} Y$ *if and only if* $P(X > Y) \geq P(Y > X)$. *When this inequality holds some authors say that* X *is preferred to* Y *in the precedence order (see [25, 26, 112]). Hence, this provides an equivalent formulation of statistical preference. We have preferred to use the latter because it provides degrees of preference between the alternatives by means of the probabilistic relation* Q. *Note that other authors consider a difference definition of precedence order ([2, 25, 26, 112, 210]) which is not equivalent in general, as we have seen in the previous remark.*

A probabilistic or reciprocal relation can also be seen as a fuzzy relation. For this reason, statistical preference can be interpreted as a defuzzyfication of the relation Q:

$$X \succeq_{SP} Y \Leftrightarrow (X, Y) \in Q_{\frac{1}{2}},$$

where $Q_{\frac{1}{2}}$ denotes the $\frac{1}{2}$-cut of Q:

$$Q_{\frac{1}{2}} = \left\{(X, Y) \in D \times D : Q(X, Y) \geq \frac{1}{2}\right\}.$$

Another connection with fuzzy set theory can be made if we consider that the information contained in the probabilistic relation can be also presented by means of a fuzzy relation. This was initially prop osed in [16, 57] and latter analyzed in detail in [122]; rece ntly, a generalization has been presented in [163]. There, from any probabilistic re lation Q defined on aset D, $h(Q)$, with h: $[0,1] \to [0,1]$ is a fuzzy weak preference relationif and only if $h\left(\frac{1}{2}\right) = 1$.

The previous resultwas proven forany probabilisticrelation Q, but whenweare comparing random variables by means of the re lation Q defined on Equation (2.7), $h(Q)$ is an order-preserving fuzzy weak preference relation if and only if $h(0) = 0$, $h(\frac{1}{2}) = 1$ and h is increasing in $[0,1]$.

The initial h prop osed in [57] was $h(x) = \min(1, 2x)$ but, of course, an infinite family of functions may be considered. Asanexample, wewillobtaintheexpressionof the weak preference relation R inthat initialcase:

$$R(X,Y) = \begin{cases} 1 & \text{if } P(X > Y) \geq P(Y > X), \\ 1 + P(X > Y) - P(Y > X) & \text{otherwise.} \end{cases}$$

Example 2.23 *Letusconsider therandomvariable X uniformlydistributedinthe interval $(4,6)$, and let Y_1, Y_2, Y_3 and Y_4 bethe uniformly distributed randomvariables in the intervals $(7,9)$, $(5,7)$, $(3,5)$ and $(0,2)$ respectively. If we assume themtobe independent, it holds that:*

$$\begin{array}{ll} Q(X,Y_1) = 0 & R(X,Y_1) = 0. \\ Q(X,Y_2) = \frac{1}{8} & R(X,Y_2) = \frac{1}{4}. \\ Q(X,Y_3) = \frac{7}{8} & R(X,Y_3) = 1. \\ Q(X,Y_4) = 1 & R(X,Y_4) = 1. \end{array}$$

We can notice the different scales used by Q and R.

Thus, weconclude that R can be seen as a “greater than or equal to” relation, but the meaning of Q is totally diffe rent. In fact, the interpretation of the value of the fuzzy relation R is: the closerthevalueto0, theweaker thepreferenceof X over Y.

We have already mentioned some advantages of statistical preference over sto chastic dominance: on the one han d, statistical preference allows the p ossibility of establishing preference degrees between the alternatives; on the otherhand statistical preference determines a total relationship between the random variables, while we can find pairs of random variables which are incomparable under the n-th degree sto chastic dominance. Another advantage is that it takes into account the p ossible dep endence between the random variables since it is based on the joint distribution, while sto chastic dominance only uses marginal distributions.

In this sense, recall that given n independent real-valued random variables X_1, ..., X_n, with cumulative distribution functions F_{X_1}, ..., F_{X_n}, respectively, the joint cumulative distribution function, denoted by F, is the product of the marginals:

$$F(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n),$$

for any $x_1, \ldots, x_n \in \mathbb{R}$. In general, the joint cumulative distribution function can be expressed by:

$$F(x_1, \ldots, x_n) = C(F_{X_1}(x_1), \ldots, F_{X_n}(x_n))$$

for any $x_1, \ldots, x_n \in \mathbb{R}$, where C is a function called copula.

Definition 2.24 ([166]) *A n-dimensional* copula *is a function $C : [0,1]^n \to [0,1]$ satisfying the following properties:*

- *For every* $(x_1, \ldots, x_n) \in [0,1]^n$, $C(x_1, \ldots, x_n) = 0$ *if* $x_i = 0$ *for some* $i \in \{1, \ldots, n\}$.
- *For every* $(x_1, \ldots, x_n) \in [0,1]^n$, $C(x_1, \ldots, x_n) = x_i$ *if* $x_j = 1$ *for every* $j \neq i$.
- *For every* $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in [0,1]^n$:

$$V_C([x, y]) \geq 0,$$

where:

$$V_C([x, y]) = \sum_{i=1}^{n} \sum_{c_i \in \{a_i, b_i\}} \operatorname{sgn}(c_1, \ldots, c_n) C(c_1, \ldots, c_n),$$

where the function sgn *is defined by:*

$$\operatorname{sgn}(c_1, \ldots, c_n) = \begin{cases} 1 & \text{if } c_i = a_i \text{ for an even number of } i\text{'s}. \\ -1 & \text{if } c_i = a_i \text{ for an odd number of } i\text{'s}. \end{cases}$$

In particular, a 2-dimensional copula (a copula, for short) is a function $C : [0,1] \times [0,1] \to [0,1]$ satisfying $C(x, 0) = C(0, x) = 0$ and $C(x, 1) = C(1, x) = x$ for every $x \in [0,1]$ and

$$C(x_1, y_1) + C(x_2, y_2) \geq C(x_1, y_2) + C(x_2, y_1)$$

for every $(x_1, x_2, y_1, y_2) \in [0,1]^4$ such that $x_1 \leq x_2$ and $y_1 \leq y_2$.

The most important examples of copulas are the following:

- The product copula π: $\pi(x_1, \ldots, x_n) = \prod_{i=1}^{n} x_i$.
- The minimum operator M: $M(x_1, \ldots, x_n) = \min\{x_1, \ldots, x_n\}$.
- The Łukasiewicz operator W, for $n = 2$: $W(x_1, x_2) = \max\{0, x_1 + x_2 - 1\}$.

Since the Łukasiewicz operator is associative, it can only be defined as n-ary operator: $W(x_1, \ldots, x_n) = \max\{0, x_1 + \ldots + x_n - (n-1)\}$. However, it is a copula only for $n=2$. One important and well-known result concerning copula is that every n-dimensional copula is bounded by the Łukasiewicz and the minimum operator:

$$W(x_1, \ldots, x_n) \leq C(x_1, \ldots, x_n) \leq M(x_1, \ldots, x_n) \text{ for every } (x_1, \ldots, x_n) \in [0,1]^n. \quad (2.8)$$

This inequality is known as the Fréchet-Hoeffding inequality. For this reason, the Łukasiewicz and the minimum operators are also called the *lower and upper Fréchet-Hoeffding bounds* ([79]).

Recall that, although W is not a copula for $n>2$, it can be approximated by a copula on each point:

Proposition 2.25 ([62, 166]) *For any $(x_1, \ldots, x_n) \in [0,1]^n$ there is a n-dimensional copula C such that $C(x_1, \ldots, x_n) = W(x_1, \ldots, x_n)$.*

In particular, when $n=2$, W is a copula and the previous result becomes trivial.

A particular type of copulas are the Archimedean copulas.

Definition 2.26 ([166]) *A n-dimensional copula C is* Archimedean *if there exists a function $\phi : [0,1] \to [0,\infty]$, called generator of C, strictly decreasing, satisfying that $-\phi$ is n-monotone, $\phi(1) = 0$ and:*

$$C(x_1, \ldots, x_n) = \phi^{[-1]}(\phi(x_1) + \ldots + \phi(x_n)), \quad (2.9)$$

for every $(x_1, \ldots, x_n) \in [0,1]^n$, where $\phi^{[-1]}$ denotes the pseudo-inverse of ϕ, and it is defined by:

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t) & \text{if } 0 \leq t \leq \phi(0). \\ 0 & \text{if } \phi(0) < t \leq \infty. \end{cases}$$

The main Archimedean copulas are the product, whose generator is $\phi_\pi(t) = -\log t$, and the Łukasiewicz operator for $n=2$, whose generator is $\phi_W(t) = 1 - t$. The most important non-Archimedean copula is the minimum operator.

Archimedean copulas can also be divided into two groups: strict and nilpotent Archimedean copulas. An Archimedean copula is called *strict* if its generator, ϕ, satisfies $\phi(0) = \infty$. In such case, the pseudo inverse becomes the inverse, and therefore Equation (2.9) becomes:

$$C(x_1, \ldots, x_n) = \phi^{-1}(\phi(x_1) + \ldots + \phi(x_n)). \quad (2.10)$$

An Archimedean copula is *nilpotent* if $\phi(0) < \infty$. The most important examples of strict and nilpotent copulas are the product and the Łukasiewicz operator, respectively.

One of the most important traits of copulas is the famous Sklar's theorem.

Theorem 2.27 ([189]) *Let $X_1, \ldots, X_n$ be n random variables, and let $F_{X_1}, \ldots, F_{X_n}$ denote their respective cumulative distribution functions. If F denotes the joint cumulative distribution function, then there exists a copula C such that*

$$F(x_1, \ldots, x_n) = C(F_{X_1}(x_1), \ldots, F_{X_n}(x_n)) \text{ for every } (x_1, \ldots, x_n) \quad \mathbb{R}^n.$$

When the copula is Archimedean, last expression becomes:

$$F(x_1, \ldots, x_n) = \phi^{-1]}(\phi(F_{X_1}(x_1)) + \ldots + \phi(F_{X_n}(x_n))).$$

Obviously, a pair of random variables is coupled by the product if and only if they are independent. Moreover, random variables coupled by the minimum operator (respectively, by the Łukasiewicz operator) are called *comonotonic* (respectively, *countermonotonic*). These two cases are very important in the theory of copulas, and for this reason we will study in detail the properties of statistical preference and stochastic dominance for them. In fact, from the Fréchet-Hoeffding bounds of Equation (2.8), an interpretation of comonotonic and countermonotonic random variables can be given. In order to see this, recall that a subset S of $\overline{\mathbb{R}}^2$ is increasing if and only if for each (x, y) $\mathbb{R}^2$ either:

1. for all (u, v) in S, $u \leq x$ implies $v \leq y$; or
2. for all (u, v) in S, $v \leq y$ implies $u \leq x$.

Similarly, a subset S of $\overline{\mathbb{R}}^2$ is decreasing if and only if for each (x, y) $\mathbb{R}^2$ either:

1. for all (u, v) in S, $u \leq x$ implies $v \geq y$; or
2. for all (u, v) in S, $v \leq y$ implies $u \geq x$.

Using this notation, the following result is presented in [166, Theorem 2.5.4] and proved in [124].

Proposition 2.28 *Let X and Y be two real-valued random variables. X and Y are comonotonic if and only if the support of the joint distribution function is a increasing subset of $\overline{\mathbb{R}}^2$, and X and Y are countermonotonic if and only if the support of the joint distribution function is a decreasing subset of $\overline{\mathbb{R}}^2$.*

When X and Y are continuous, we say that Y is almost surely an increasing function of X if and only if X and Y are comonotonic, and Y is almost surely a decreasing function of X if and only if they are countermonotonic.

2.2 Imprecise probabilities

Next, we discuss briefly *imprecise probability models*. This is the generic term used to refer to all mathematical models that serve as an alternative and a generalization of probability models to situations where our knowledge if vague or scarce. It includes possibility measures ([217]), Choquet capacities ([39]), belief functions ([187]) or coherent lower previsions ([205]), among other models.

2.2.1 Coherent lower previsions

We begin by introducing the main concepts of the theory of coherent lower previsions. Consider a possibility space Ω. A *gamble* is a real-valued functional defined on Ω. We shall denote by $\mathcal{L}(\Omega)$ the set of all gambles on Ω, while $\mathcal{L}^+(\Omega)$ denotes the set of positive gambles on Ω. Given a subset A of Ω, the indicator function of A is the gamble that takes the value 1 on the elements of A and 0 elsewhere. We shall denote this gamble by I_A, or by A when no confusion is possible.

A *lower prevision* is a functional $\underline{P}$ defined on a set of gambles $\mathcal{K} \subseteq \mathcal{L}(\Omega)$. Given a gamble f, $\underline{P}(f)$ is understood to represent a subject's supremum acceptable buying price for f, in the sense that for any $\varepsilon > 0$ the transaction $f - \underline{P}(f) + \varepsilon$ is acceptable to him.

Using this interpretation, we can derive the notion of coherence.

Definition 2.29 ([205, Section 2.5]) *Consider the lower prevision $\underline{P}: \mathcal{K} \to \mathbb{R}$, where $\mathcal{K} \subseteq \mathcal{P}(\Omega)$. It* avoids sure loss *if for any natural number n and any $f_1, \dots, f_n \in \mathcal{K}$ it holds that:*

$$\sup_{\omega \in \Omega} \sum_{k=1}^{n} [f_k(\omega) - \underline{P}(f_k)] \geq 0.$$

Also, $\underline{P}$ is coherent *iff for any natural numbers n and m and $f_0, f_1, \dots, f_n \in \mathcal{K}$, it holds that:*

$$\sup_{\omega \in \Omega} \sum_{i=1}^{n} [f_k(\omega) - \underline{P}(f_k)] - m[f_0(\omega) - \underline{P}(f_0)] \geq 0.$$

The interpretation of this notion is that the acceptable buying prices encompassed by $\{\underline{P}(f) : f \in \mathcal{L}(\Omega)\}$ are consistent with each other, in the sense defined in [205, Section 2.5]. From any lower prevision $\underline{P}$ it is possible to define a set of probabilities, also called credal set, by:

$$\mathcal{M}(\underline{P}) = \{P \text{ finitely additive probabilities} : P \geq \underline{P}\}.$$

The following result relates coherence and avoiding sure loss to the credal set $\mathcal{M}(\underline{P})$. It is usually called the *Envelope Theorem*.

Theorem 2.30 ([205, Section 3.3.3]) *Let $\underline{P}$ be a lower probability defined on a set of gambles $\mathcal{K}$, and let $\mathcal{M}(\underline{P})$ denote its associated credal set. Then:*

$$\underline{P} \text{ avoids sure loss} \Leftrightarrow \mathcal{M}(\underline{P}) \neq \emptyset$$

and

$$\underline{P} \text{ is coherent} \Leftrightarrow \underline{P}(f) = \inf_{P \in \mathcal{M}(\underline{P})} P(f).$$

By conjugacy, an operator $\overline{P}$ defined on a set of gambles $\mathcal{K}$ is called *upper prevision*. For any $f \in \mathcal{K}$, $\overline{P}(f)$ is understood to represent the subject's infimum acceptable selling price for f, in the sense that for any $\varepsilon > 0$ the transaction $\overline{P}(f) + \varepsilon - f$ is acceptable to him. An upper prevision avoids sure loss (respectively, is coherent) if and only if $\overline{P}(f) = -\underline{P}(-f)$, where $\underline{P}$ is a lower prevision that avoids sure loss (respectively, that is coherent).

When the domain $\mathcal{K}$ of the lower and upper previsions is formed by subsets of Ω, $\underline{P}$ and $\overline{P}$ are called *lower and upper probabilities*, respectively.

Next proposition shows several properties of coherent lower and upper probabilities.

Proposition 2.31 ([205, Section 2.4.7]) *Let $\underline{P}$ be a lower probability and let $\overline{P}$ denote its conjugate upper probability. The following statements hold for any $A, B \subseteq \Omega$:*

$$A \cap B = \emptyset \Rightarrow \underline{P}(A \cup B) \geq \underline{P}(A) + \underline{P}(B). \tag{2.11}$$

$$A \cap B = \emptyset \Rightarrow \overline{P}(A \cup B) \geq \underline{P}(A) + \overline{P}(B). \tag{2.12}$$

$$\underline{P}(A) + \underline{P}(B) \leq \underline{P}(A \cup B) + \overline{P}(A \cap B). \tag{2.13}$$

$$\overline{P}(A \cup B) + \underline{P}(A \cap B) \geq \underline{P}(A) + \underline{P}(B). \tag{2.14}$$

$$\overline{P}(A \cup B) + \overline{P}(A \cap B) \geq \overline{P}(A) + \underline{P}(B). \tag{2.15}$$

Given a coherent lower prevision $\underline{P}$ with domain $\mathcal{K}$, we may be interested in extending $\underline{P}$ to a more general domain $\mathcal{K}' \supseteq \mathcal{K}$. This can be made by means of the natural extension.

Definition 2.32 ([205, Section 3.1]) *Let $\underline{P}$ be a coherent lower prevision on $\mathcal{K}$, and consider $\mathcal{K} \subseteq \mathcal{K}'$. Then, for any $f \in \mathcal{K}'$, the* natural extension *of $\underline{P}$ is defined by:*

$$\underline{E}(f) = \inf_{P \in \mathcal{M}(\underline{P})} P(f).$$

The natural extension is the least committal, that is the most imprecise, coherent extension of $\underline{P}$.

One instance where coherent lower previsions appear is when dealing with p-boxes.

Definition 2.33 ([75]) A probability box, *or p-box for short,* $(\underline{F}, \overline{F})$ *is the set of cumulative distribution functions bounded between two finitely additive distribution functions* $\underline{F}$ *and* $\overline{F}$ *such that* $\underline{F} \leq \overline{F}$. *We shall refer to* $\underline{F}$ *as the* lower distribution function *and to* $\overline{F}$ *as the* upper distribution function *of the p-box.*

Note that $\underline{F}, \overline{F}$ need not be cumulative distribution functions, and as such they need not belong to the set $(\underline{F}, \overline{F})$; they are only required to be finitely additive distribution functions. In particular, if we consider a set $\mathcal{F}$ of distribution functions, its associated *lower* and *upper* distribution functions are given by

$$\underline{F}(x) := \inf_{F \in \mathcal{F}} F(x), \quad \overline{F}(x) := \sup_{F \in \mathcal{F}} F(x). \tag{2.16}$$

Proposition 2.34 *Given a set of cumulative distribution functions* $\mathcal{F}$, *its lower bound* $\underline{F}$ *is also a cumulative distribution function, while* $\overline{F}$ *is a finitely additive cumulative distribution function.*

P-boxes have been connected to info-gap theory ([76]), random sets ([103, 172]), and possibility measures ([17, 51, 198]).

Given a p-box $(\underline{F}, \overline{F})$ on Ω, it induces a lower probability $\underline{P}_{(\underline{F},\overline{F})}$ on the set

$$\mathcal{K} = \{A_x, A_x^c : x \in \Omega\},$$

where $A_x = \{x' \in \Omega : x' \leq x\}$, by:

$$\underline{P}_{(\underline{F},\overline{F})}(A_x) = \underline{F}(x) \text{ and } \underline{P}_{(\underline{F},\overline{F})}(A_x^c) = 1 - \overline{F}(x). \tag{2.17}$$

If $\underline{F} = \overline{F} = F$, $\underline{P}_{(\underline{F},\overline{F})}$ is usually denoted by P_F. The following result is stated in [209] and proved in [198, 201].

Theorem 2.35 ([198, Section 3], [201, Theorem 3.59]) *Consider two maps* $\underline{F}$ *and* $\overline{F}$ *from* Ω *to* $[0,1]$ *and let* $\underline{P}_{(\underline{F},\overline{F})} : \mathcal{K} \to [0,1]$ *be the lower probability they induce by means of Equation* (2.17). *The following statements are equivalent:*

- $\underline{P}_{(\underline{F},\overline{F})}$ *is a coherent lower probability.*
- $\underline{F}, \overline{F}$ *are distribution functions and* $\underline{F} \leq \overline{F}$.
- $P_{\underline{F}}$ *and* $P_{\overline{F}}$ *are coherent and* $\underline{F} \leq \overline{F}$.

In particular, if $\underline{F} = \overline{F} = F$, *then* P_F *is coherent if and only if* F *is a distribution function.*

A particular case appears when defining coherent lower previsions in product spaces $\Omega_1 \times \Omega_2$. If $\underline{P}$ is a coherent lower prevision taking values on $\mathcal{L}(\Omega_1 \times \Omega_2)$, we can consider its marginals $\underline{P}_1$ or $\underline{P}_2$ as coherent lower previsions on $\mathcal{L}(\Omega_1)$ or $\mathcal{L}(\Omega_2)$, respectively, defined by:

$$\underline{P}_1(f) = \underline{P}(f) \quad \text{and} \quad \underline{P}_2(f) = \underline{P}(f)$$

for any gamble f on $\Omega_1 \times \Omega_2$. They will arise when trying to define coherent lower previsions from bivariate p-boxes.

In this work, we shall use imprecise probability models because we shall be interested in the comparison of sets of alternatives, each with its associated probability distribution; we obtain thus a set $\mathcal{P}$ of probability measures. This set can be summarized by means of its *lower* and *upper* envelopes, which are given by:

$$\underline{P}(A) := \inf_{P \in \mathcal{P}} P(A), \; \overline{P}(A) := \sup_{P \in \mathcal{P}} P(A), \tag{2.18}$$

and which are *coherent* lower and upper probabilities.

2.2.2 Conditional lower previsions

Consider two random variables X and Y taking values in two spaces Ω_1 and Ω_2 and let $\underline{P}$ be a coherent lower prevision taking values on $\mathcal{L}(\Omega_1 \times \Omega_2)$. We define a conditional lower prevision $\underline{P}(\cdot \mid Y)$ as a function with two arguments. For any $y \in \Omega_2$, $\underline{P}(\cdot \mid y)$ is real functional on the set $\mathcal{L}(\Omega_1 \times \Omega_2)$, while for any gamble f on $\Omega_1 \times \Omega_2$, $\underline{P}(f \mid y)$ is the lower prevision of f, conditional on $\Omega_2 = y$. $\underline{P}(f \mid Y)$ is then the gamble on Ω_1 that assumes the value $\underline{P}(f \mid y)$ in y. Similar considerations can be made for $\underline{P}(\cdot \mid X)$.

Definition 2.36 *The conditional lower prevision $\underline{P}(\cdot \mid Y)$ is called* separately coherent *if for all $y \in \Omega_2$, $\lambda \geq 0$ and $f, g \in \mathcal{L}(\Omega_1 \times \Omega_2)$ it satisfies the following conditions:*

SC1 $\underline{P}(f \mid y) \geq \inf_{x \in \Omega} f(x, y)$.
SC2 $\underline{P}(\lambda f \mid y) = \lambda \underline{P}(f \mid y)$.
SC3 $\underline{P}(f + g \mid y) \geq \underline{P}(f \mid y) + \underline{P}(g \mid y)$.

It is known that from separate coherence the following properties hold (see [205, Theorems 6.2.4 and 6.2.6]):

$$\underline{P}(g \mid y) = \underline{P}(g(\cdot, y) \mid y) \text{ and } \underline{P}(fg \mid Y) = f\underline{P}(g \mid Y),$$

for all $y \in \Omega_2$, all positive gambles f on Ω_2 and all gambles g on $\Omega_1 \times \Omega_2$.

We now investigate separate coherence and coherence together. For any gamble f on $\mathcal{L}(\Omega_1 \times \Omega_2)$, we define:

$$G(f \mid y) = I_{\{y\}}[f - \underline{P}(f \mid Y)] = I_{\{y\}}[f(\cdot, y) - \underline{P}(f(\cdot, y) \mid y)]$$

and

$$G(f \mid Y) = f - \underline{P}(f \mid Y) = f - \underline{P}(f \mid Y) = \sum_{y \ \Omega} I_{\{y\}}[f(\ , y) - \underline{P}(f(\ , y) \mid y)].$$

Definition 2.37 *Let $\underline{P}(\ \mid Y)$ and $\underline{P}(\ \mid X)$ be two separately coherent conditional lower previsions. They are called* weakly coherent *if and only if for all $f_1, f_2 \ \mathcal{L}(\Omega_1 \times \Omega_2)$, all $x \ \Omega_1, y \ \Omega_2$ and $g \ \mathcal{L}(\Omega_1 \times \Omega_2)$, there are some*

$$B_1 \ supp_{\Omega_1}(f_2) \ supp_{\Omega_2}(f_1) \ (\{x\} \times \Omega_2)$$
$$B_2 \ supp_{\Omega_1}(f_2) \ supp_{\Omega_2}(f_1) \ (\Omega_1 \times \{y\})$$

such that:

$$\sup_{z \ B_1} [G(f_1 \mid Y) + G(f_2 \mid X) - G(g \mid x)](z) \geq 0$$

and

$$\sup_{z \ B_2} [G(f_1 \mid Y) + G(f_2 \mid X) - G(g \mid y)](z) \geq 0,$$

where

$$supp_{\Omega_1}(f) = \{\{x\} \times \Omega_2, x \ \Omega_1 \mid f(x, \ ) = 0\}$$

and

$$supp_{\Omega_2}(f) = \{\Omega_1 \times \{y\}, y \ \Omega_2 \mid f(\ , y) = 0\}.$$

We say that $\underline{P}(\ \mid Y)$ and $\underline{P}(\ \mid X)$ are coherent *if for all $f_1, f_2 \ \mathcal{L}(\Omega_1 \times \Omega_2)$, all $x \ \Omega_1, y \ \Omega_2$ and all $g \ \mathcal{L}(\Omega_1 \times \Omega_2)$ it holds that:*

$$\sup_{z \ \Omega_1 \times \Omega_2} [G(f_1 \mid Y) + G(f_2 \mid X) - G(g \mid x)](z) \geq 0.$$
$$\sup_{z \ \Omega_1 \times \Omega_2} [G(f_1 \mid Y) + G(f_2 \mid X) - G(g \mid y)](z) \geq 0.$$

Several res ults can b e found in the literature relating coherence and weak coherence.

Theorem 2.38 ([137, Theorem 1]) *Let $\underline{P}(\ \mid X)$ and $\underline{P}(\ \mid Y)$ be separately coherent conditional lower previsions. Theyareweaklycoherent if and only if there is some coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega_1 \times \Omega_2)$ such that*

$$\underline{P}(G(f \mid X)) \geq 0 \text{ and } \underline{P}(G(f \mid x)) = 0 \quad \text{for any } f \ \mathcal{L}(\Omega_2), x \ \Omega_2,$$
$$\underline{P}(G(g \mid Y)) \geq 0 \text{ and } \underline{P}(G(g \mid y)) = 0 \quad \text{for any } g \ \mathcal{L}(\Omega_1), y \ \Omega_1.$$

The followingresult isknownasthe ReductionTheorem.

Theorem 2.39 ([205, Theorem 7.1.5]) *Let $\underline{P}(\ \mid X)$ and $\underline{P}(\ \mid Y)$ be separately coherent conditional lower previsions defined on $\mathcal{L}(\Omega_1 \times \Omega_2)$, and let $\underline{P}$ bea coherent lower prevision on $\mathcal{L}(\Omega_1 \times \Omega_2)$. Then $\underline{P}$, $\underline{P}(\ \mid X)$ and $\underline{P}(\ \mid Y)$ are coherent if and only if the fol lowing two conditions holds:*

1. *$\underline{P}$, $\underline{P}(\ \mid X)$ and $\underline{P}(\ \mid Y)$ are weakly coherent.*
2. *$\underline{P}(\ \mid X)$ and $\underline{P}(\ \mid Y)$ are coherent.*

2.2.3 Non-additive measures

One important example of coherent lower previsions are the n-monotone ones, which were first introduced by Choquet in [39].

Definition 2.40 ([39]) *A coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega)$ is called n-monotone if and only if:*

$$\underline{P}\left(\bigvee_{i=1}^{p} f_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1,\dots,p\}} (-1)^{|I|+1}\, \underline{P}\left(\bigwedge_{i \in I} f_i\right)$$

for all $2 \leq p \leq n$ and all $f_1, \dots, f_p$ in $\mathcal{L}(\Omega)$, where $\vee$ denotes the point-wise maximum and $\wedge$ the point-wise minimum.

In particular, a coherent lower probability $\underline{P} : \mathcal{P}(\Omega) \to [0,1]$ is n-monotone when

$$\underline{P}\left(\bigcup_{i=1}^{p} A_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1,\dots,p\}} (-1)^{|I|+1}\, \underline{P}\left(\bigcap_{i \in I} A_i\right)$$

for all $2 \leq p \leq n$ and all subsets $A_1, \dots, A_p$ of Ω.

A coherent lower prevision on $\mathcal{L}(\Omega)$, that is n-monotone for all $n \in \mathbb{N}$, is called *completely monotone*, and its restriction to events is a *belief function*. The restriction to events of the conjugate upper prevision is called *plausibility function*. Belief and plausibility functions are usually denoted by bel and pl.

Another type of non-additive measure are possibility measures.

Definition 2.41 ([70]) A possibility measure *on [0,1] is a supremum preserving set function* $\Pi : \mathcal{P}([0,1]) \to [0,1]$. *It is characterised by its restriction to events π, which is called its* possibility distribution. *The conjugate function N of a possibility measure is called a* necessity measure:

$$N(A) = 1 - \Pi(A^c).$$

Because of their computational simplicity, possibility measures are widely applied in many fields, including data analysis ([196]), diagnosis ([33]), cased-based reasoning ([91]) and psychology ([177]).

Let us see how to apply our extension stochastic dominance to the comparison of possibility measures; another approach to preference modeling with possibility measures is discussed in [19, 115].

The connection between possibility measures and p-boxes was already explored in [199], and it was proven that almost any possibility measure can be seen as the natural extension of a corresponding p-box. However, the definition of this p-box implies defining some particular order on our referential space, which could be different to the one we already have there (for instance if the possibility measure is defined on [0,1] it may seem counterintuitive to consider anything different from the natural order), and moreover two different possibility measures may produce two different orders on the same space, making it impossible to compare them.

Instead, we shall consider a possibility measure Π on $\Omega = [0,1]$, its associated set of probability measures:

$$\mathcal{M}(\Pi) := \{P \text{ probability} : P(A) \leq \Pi(A) \ \forall A\}, \tag{2.19}$$

and the corresponding set of distribution functions $\mathcal{F}$. Let $(\underline{F}, \overline{F})$ be its associated p-box.

Since any possibility measure on [0,1] can be obtained as the upper probability of a random set ([84]), and moreover in that case ([131]) the upper probability of the random set is the maximum of the probability distributions of the measurable selections, we deduce that the p-box associated to $\mathcal{F}$ is determined by the following lower and upper distribution functions:

$$\overline{F}(x) = \sup_{P \leq \Pi} P([0,x]) = \Pi([0,x]) = \sup_{y \leq x} \pi(y) \tag{2.20}$$

$$\underline{F}(x) = \inf_{P \leq \Pi} P([0,x]) = 1 - \Pi((x,1]) = 1 - \sup_{y > x} \pi(y).$$

Note however, that these lower and upper distribution functions need not belong to $\mathcal{F}$: if for instance we consider the possibility measure associated to the possibility distribution $\pi = I_{(0.5,1]}$, we obtain $\underline{F} = \pi$, which is not right-continuous, and consequently cannot belong to the set $\mathcal{F}$ of distribution functions associated to $\mathcal{M}(\Pi)$.

Another interesting type of non-additivity measures, that includes possibility measures as a particular case are clouds. Following Neumaier ([168]), a *cloud* is a pair of functions $[\delta, \pi]$ where $\pi, \delta : [0,1] \to [0,1]$ satisfy:

- $\delta \leq \pi$.
- There exists $x \in [0,1]$ such that $\pi(x) = 0$.
- There exists $y \in [0,1]$ such that $\delta(y) = 1$.

δ and π are called the *lower* and *upper distributions* of the cloud, respectively.

Any cloud $[\delta, \pi]$ has an associated set of probabilities $\mathcal{P}_{[\delta,\pi]}$, that is the set of probabilities P satisfying:

$$P(\{x \in [0,1] \mid \delta(x) \geq \alpha\}) \leq 1 - \alpha \leq P(\{x \in [0,1] \mid \pi(x) > \alpha\}).$$

Since both π and $1-\delta$ are possibility distributions we can consider their asso ciated credal sets P_π and $P_{1-\delta}$, given by

$$P_\pi := \{P \text{ probability } : P(A) \leq \Pi(A) \quad A \quad \beta_{[0,1]}\},$$

where Π denotes the p ossibility measure asso ciated to the possibility distribution π, and similarly for $P_{1-\delta}$. From [65], it holds that $P_{[\delta,\pi]} = P_{1-\delta} \cap P_\pi$.

2.2.4 Random sets

One context where completel y monotone lower previsions arise naturally is that of measurable multi-valued mappings, or randomsets ([59,96]).

Definition 2.42 *Let* $(\Omega, \mathcal{A}, P)$ *be aprobability space,* $(\Omega, \mathcal{A})$ *ameasurable space,and* $\Gamma : \Omega \to \mathcal{P}(\Omega)$ *a non-empty multi-valued mapping. It is cal led* random set *when*

$$\Gamma(A) = \{\omega \quad \Omega : \Gamma(\omega) \cap A = \quad\} \quad A$$

for any $A \quad \mathcal{A}$.

One instance of random sets are random intervals, that are those satisfying that $\Gamma(\omega)$ is an interval for any $\omega \quad \Omega$.

If Γ mo dels the imprecise knowledge ab out a random variable X, $\Gamma(\omega)$ represents that the “true” value of $X(\omega)$ belongs to $\Gamma(\omega)$. Then, all we know ab out X isthat itis oneof the *measurable selections* of Ω:

$$S(\Gamma) = \{U : \Omega \to \Omega \text{ random variable } : U(\omega) \quad \Gamma(\omega) \quad \omega \quad \Omega\}. \tag{2.21}$$

This interpretation of multi-valued mappings as a mo del for the imprecise knowledge of a random variable is not new, and can b e traced back to Krus e and Meyer ([104]).The *epistemic* interpretation contrasts withthe *ontic* interpretation which issometimes given to random sets as naturally impreci se quantities ([73]).

Random sets generate upp er and lower probabilities.

Definition 2.43 ([59]) *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space* $(\Omega, \mathcal{A})$ *a measurable space and* $\Gamma : \Omega \to \mathcal{P}(\Omega)$ *arandom set. Then its* upper *and* lower probabilities *are the functions* $P, P : \mathcal{A} \to [0,1]$ *given by:*

$$P(A) = P(\{\omega : \Gamma(\omega) \cap A = \quad\}) \text{ and } P(A) = P(\{\omega : \quad = \Gamma(\omega) \quad A\}) \tag{2.22}$$

for any $A \quad \mathcal{A}$ *. These upperand lowerprobabilities are,in particular, a plausibility and a belief function, respectively. Furthermore,they define thecredal set* $M(P_\Gamma)$ *given by:*

$$M(P_\Gamma) = \{P \text{ probability } : P_\Gamma(A) \leq P(A) \leq P_\Gamma(A) \quad A \quad \mathcal{A}\}. \tag{2.23}$$

The upp er and lower probabilities of a random set are in particular coherent lower and upp er probabilities, and constitute the lower and upp er bounds of the probabilities induced bythemeasurable selections:

$$P_*(A) \leq P_X(A) \leq P^*(A) \text{ for every } X \in S(\Gamma). \quad (2.24)$$

Therefore, their asso ciated cumulative distribution functions provide lower and upp er b ounds of the lower and upper distribution functions asso ciated to $S(\Gamma)$. The inequalities of Equation (2.24) can b e strict [130, Example 1]; howeve r, under fairly general conditions

$$P^*(A) = \max P(\Gamma)(A) \text{ and } P_*(A) = \min P(\Gamma)(A) \text{ for every } A \in \mathcal{A}, \quad (2.25)$$

where $P(\Gamma)(A) = \{P_X(A) : X \in S(\Gamma)\}$. Inparticular, if Γ takes values on the measurable space $([0,1], \beta_{[0,1]})$, wh ere $\beta_{[0,1]}$ denotes theBorel σ-field, Equation (2.25) hol ds under any of the following cond itions ([130]):

- Ifthe class $\{\Gamma(\omega) : \omega \in \Omega\}$ is countable.
- If $\Gamma(\omega)$ is closed for every $\omega \in \Omega$.
- If $\Gamma(\omega)$ is op en for every $\omega \in \Omega$.

However, thetwo setsarenot equivalentingeneral, and $M(P_\Gamma)$ can only b e see n as an outer approxim ation. There are nonetheless situations in which both sets coincide. First, let us intro duce the following definition.

Definition 2.44 *Consider twofunctions $A, B : \Omega \to \mathbb{R}$. They are cal led strictly comonotone if $(A(\omega) - A(\omega'))\geq 0$ if and only if $(B(\omega) - B(\omega')) \geq 0$ for any $\omega, \omega' \in \Omega$.*

A similar but less restrictive notionisthe one of comonotone functions: A and B are called comonotone if $(A(\omega) - A(\omega'))(B(\omega) - B(\omega')) \geq 0$ for any $\omega, \omega' \in \Omega$. Note that both notions are not equivalentin general. Infact, two increasingandcomonotonefunctions A and B arestrictly comonotoneif andonly if $A(\omega') = A(\omega)$ if and only if $B(\omega') = B(\omega)$, and two comonotone functions A and B with $A=0 \leq B$ are strictly comonotone if and only if B is constant.

Next,welist some situationsin which the sets $P(\Gamma)$ and $M(P_\Gamma)$ coincide.

Prop osition 2.45 ([129]) *Let $(\Omega, \mathcal{A}, P)$ beaprobabilityspace and consider the random closed interval $\Gamma := [A, B] : \Omega \to \mathcal{P}(\mathbb{R})$. Let $P(\Gamma), M(P_\Gamma)$ denote the sets of probability measures induced by the selections and those dominated by the upper probability, respectively. Then:*

1. $P_\Gamma(C) = \max\{Q(C) : Q \in P(\Gamma)\} \; \forall C \in \beta_{\mathbb{R}}$.

2. *$M(P_\Gamma) = \overline{Conv(P(\Gamma))}$, and if $(\Omega, \mathcal{A}, P)$ is non-atomic then $M(P_\Gamma) = \overline{P(\Gamma)}$.*
3. *When $(\Omega, \mathcal{A}, P) = ([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$, the equality $M(P_\Gamma) = P(\Gamma)$ holds under any of the fol lowing conditions:*
   - *(a) The variables $A, B : [0,1] \to \mathbb{R}$ are increasing.*
   - *(b) $A=0 \leq B$.*
   - *(c) A,B are strictlycomonotone.*

For a complete study on the conditions under which the lower and upp er probabilities are attained or the conditions under whichthe sets $P(\Gamma)$ and $M(P)$ coincide, we refer to [125].

Theorem 2.46 ([130, Theorem 14]) *Let $(\Omega, \mathcal{A}, P)$ be aprobability space. Consider the measurable space $([0,1], \beta_{[0,1]})$ and let $\Gamma : \Omega \to P([0,1])$ be arandom set. If $P(A) = \max P(\Gamma)(A)$ for all $A \quad \mathcal{A}$, then for any bounded random variable $f : [0,1] \to \mathbb{R}$:*

$$(C) \quad f dP = \sup_{U \quad S(\Gamma)} \quad f dP_U, \qquad (C) \quad f dP = \inf_{U \quad S(\Gamma)} \quad f dP_U,$$

and consequently:

$$(C) \quad f dP = \sup(A) \quad (f \circ \Gamma) dP, \qquad (C) \quad f dP = \inf(A) \quad (f \circ \Gamma) dP,$$

where $(C) \quad f dP$ denotesthe Choquet integraloff with respectto P, and $(A) \quad (f \circ \Gamma) dP$ denotes the Aumann int egralof $f \circ \Gamma$ with respect to P, given by:

$$(A) \quad (f \circ \Gamma) dP = \quad f dP_U : U \quad S(\Gamma) \quad . \tag{2.26}$$

The upp er probability induced by a random set is always completely alternating and lower continuous [169]. Undersomeadditionalconditions, itisinparticularmaxitiveor a possibility measure:

Prop osition 2.47 ([128, Corollary 5.4]) *Let $(\Omega, \mathcal{A}, P)$ be a probability spaceand consider the random closedinterval $\Gamma : \Omega \to P(\mathbb{R})$. The fol lowing are equivalent:*

- *(a) P_Γ is a possibility measure.*
- *(b) P_Γ is maxitive.*
- *(c) Thereexists some $N \quad \Omega$ null such that for every $\omega_1, \omega_2 \quad \Omega \setminus N$, either $\Gamma(\omega_1) \quad \Gamma(\omega_2)$ or $\Gamma(\omega_2) \quad \Gamma(\omega_1)$.*

See also [50] forrelated results when $\Omega = [0,1]$.

2.3 Intuitionistic fuzzy sets

Fuzzy sets were introduced by Zadeh ([214]) as a suitable model for situations where crisp sets did not convey appropriately the available information. However, there are also situations were a more general model than fuzzy sets is deemed adequate.

A fuzzy set A assigns to every point on the universe a number $[0, 1]$ that measures the degree in which this point is compatible with the characteristic described by A. Thus, if $A(\omega)$ denotes the membership degree of ω to A, $1 - A(\omega)$ stands for the degree in which ω does not belong to A. However, two problems can arise in this situation:

1. $1 - A(\omega)$ could include at the same time both the degree of non-membership and the degree of uncertainty or indeterminacy.
2. The membership degree could not be precisely described.

Consider the following example for the former case:

Example 2.48 *Let A be the set $A =$ "objects possessing some characteristic". Thus, $A(\omega)$ stands for the degree in with ω is in accord with the given characteristic, and $1 - A(\omega)$ is the degree in which ω is not. However, ω could be partly indifferent to the characteristic. To deal with this situation, we can denote by $\mu_A(\omega) = A(\omega)$ the membership degree of ω in A, and let us define by $\nu_A(\omega)$ the degree in which ω does not belong A. Such sets, where a membership and non-membership degree is associated with any element, are called (Atanassov) Intuitionistic Fuzzy Sets (in short, IF-sets). A good example of these situations is voting, since human voters can be grouped in three classes:* vote for, vote against *or* abstain *([195]).*

In order to illustrate second scenario, consider the following example:

Example 2.49 *We are studying some element with melting temperature is m and vaporization temperature is v (obviously, $m \leq v$). For example, for water $m = 0$ ^oC and $v = 100$ ^oC. If the element is in a liquid state, we know that its temperature is greater than m, because otherwise it would be solid, and smaller than v, because otherwise it would be in gaseous state. Then, although we cannot state the exact temperature of the element, we can say for sure that it belongs to the interval $[m, v]$.*

If $A(\omega)$ denotes the (non-precisely known) membership degree of ω to A, we can consider an interval $[l_A(\omega), u_A(\omega)]$ that represents that the exact membership degree of ω to A belongs to such interval. These sets, where any element has an associated interval that bounds of the membership degree of the element to the set, are called Interval Valued Fuzzy Sets (IVF-sets, for short).

In this section we intro duce the definition and main prop erties of b oth IF-sets and IVF-sets, andweseehowthe usual op erations b etween crisp sets can b e generalized into this context. In particular, we show that b oth kind of sets are formally equivalent although, as we have already mentioned, their philosophy is different.

Let us begin with the formal definition of an intuitionistic fuzzy set.

Definition 2.50 ([4]) *Let Ω be auniverse. An* intuitionisticfuzzyset *is defined by:*

$$A = \{(\omega, \mu_A(\omega), \nu_A(\omega)) \mid \omega \in \Omega\},$$

where μ_A and ν_A are functions:

$$\mu_A, \nu_A : \Omega \to [0, 1]$$

satisfying $\mu_A(\omega) + \nu_A(\omega) \leq 1$. The function $\pi_A(\omega) = 1 - \mu_A(\omega) - \nu_A(\omega)$ is cal led the hesitation index *and it expresses the lack of know ledge on the membership of ω to A.*

We shalldenote the set of all IF-sets on Ω by $IFSs(\Omega)$.

When A is afuzzy set, its complementaryis given by $A^c = 1 - A$. That is, the memb ership degree ofevery element to the complementary of A is one minus the memb ership degree to A. Then, everyfuzzy set isin particularan IF-set wherethe hesitation index equals zero. If $FS(\Omega)$ denotes all fuzzy sets on Ω, $FS(\Omega) \subseteq IFSs(\Omega)$. For prop er IF-sets, if μ_A and ν_A denote the memb ership and non-memb ership functions, the complementary of A is defined by:

$$A^c = \{(\omega, \nu_A(\omega), \mu_A(\omega)) \mid \omega \in \Omega\}.$$

Recall that, since the emptyset $\emptyset$ is the set with no elem ents, it can b e also seen as an IF-set give n by:

$$\emptyset = \{(\omega, 0, 1) \mid \omega \in \Omega\}.$$

Similarly, full p ossi bility space Ω is the set that includes all the elements, and the re fore it can be seen as an IF-set given by:

$$\Omega = \{(\omega, 1, 0) \mid \omega \in \Omega\}.$$

Definition 2.51 ([6]) *An* intervalvaluedfuzzyset *is defined by:*

$$A = \{[l_A(\omega), u_A(\omega)] : \omega \in \Omega\},$$

where $0 \leq l_A \leq u_A(\omega) \leq 1$. When $l_A(\omega) = u_A(\omega)$ for any $\omega \in \Omega$, A becomesa fuzzy set with membership function l_A.

If $[l_A(\omega), u_A(\omega)]$ *represents that the exact membership degree of* ω *to* A *belongs to this interval, the interval* $[1 - u_A(\omega), 1 - l_A(\omega)]$ *tells us that the exact membership degree of* ω *to* A^c *belongs to such interval. Then,* A^c *is defined by:*

$$A^c = \{[1 - u_A(\omega), 1 - l_A(\omega)] : \omega \in \Omega\}.$$

Moreover, the empty set is defined by the interval $[0, 0]$ *for any* $\omega \in \Omega$*, and the total set is defined by the interval* $[1, 1]$ *for any* $\omega \in \Omega$*.*

IF-sets and IVF-sets are formally equivalent. On the one hand, given an IF-set A with membership and non-membership functions μ_A and ν_A, it defined an IVF-set by:

$$\{[\mu_A(\omega), 1 - \nu_A(\omega)] : \omega \in \Omega\}.$$

On the other hand, given an IVF-set with lower and upper bounds l_A and u_A, it defines an IF-set by:

$$\{(\omega, l_A(\omega), 1 - u_A(\omega)) : \omega \in \Omega\}.$$

For this reason, although the remainder of this section is written in terms of IF-sets, it could be analogously be formulated in terms of IVF-sets.

Let us see how to extend the usual definitions between fuzzy sets, like intersections, unions or differences, towards IF-sets. Similarly to the fuzzy case, unions and intersections of IF-sets are defined by means of t-conorms and t-norms. Recall that a t-norm is a commutative, monotonic and associative binary operator from $[0, 1] \times [0, 1]$ to $[0, 1]$ with neutral element 1, while a t-conorm satisfies the same properties than a t-norm but its neutral element is 0. From a t-norm T it is possible to define a t-conorm S_T, called the dual t-conorm, by:

$$S_T(x, y) = 1 - T(1 - x, 1 - y) \text{ for any } (x, y) \in [0, 1]^2.$$

See [99] for a complete study on t-norms.

Definition 2.52 ([63]) *Let* A *and* B *be two IF-sets given by:*

$$A = \{(\omega, \mu_A(\omega), \nu_A(\omega)) \mid \omega \in \Omega\}.$$
$$B = \{(\omega, \mu_B(\omega), \nu_B(\omega)) \mid \omega \in \Omega\}.$$

Let T *be a t-norm and* S_T *its dual t-conorm.*

- *The* T*-intersection of* A *and* B *is the IF-set* $A \cap_T B$ *defined by:*

$$A \cap_T B = \{(\omega, T(\mu_A(\omega), \mu_B(\omega)), S_T(\nu_A(\omega), \nu_B(\omega))) \mid \omega \in \Omega\}.$$

- *The* S_T*-union of* A *and* B *is the IF-set* $A \cup_{S_T} B$ *given by:*

$$A \cup_{S_T} B = \{(\omega, S_T(\mu_A(\omega), \mu_B(\omega)), T(\nu_A(\omega), \nu_B(\omega))) \mid \omega \in \Omega\}.$$

Recall that we shall use the minimum, T_N, andthe maximum, S_{T_M}, inorder tomake intersections and unions, resp ectively,since they are th e most usual op erators used in the literature. Inthatcase, the T-intersection and the S_T-union become:

$$
\begin{aligned}
A \cap_{T_M} B &= \{(\omega, T_M(\mu_A(\omega), \mu_B(\omega)), S_M(\nu_A(\omega), \nu_B(\omega))) \mid \omega \quad \Omega\} \\
&= \{(\omega, \min(\mu_A(\omega), \mu_B(\omega)), \max(\nu_A(\omega), \nu_B(\omega))) \mid \omega \quad \Omega\}. \\
A \quad {}_{S_{T_M}} B &= \{(\omega, S_M(\mu_A(\omega), \mu_B(\omega)), T_M(\nu_A(\omega), \nu_B(\omega))) \mid \omega \quad \Omega\} \\
&= \{(\omega, \max(\mu_A(\omega), \mu_B(\omega)), \min(\nu_A(\omega), \nu_B(\omega))) \mid \omega \quad \Omega\}.
\end{aligned}
$$

For simplicity,we shall denote the T-intersection and the S_T by $\cap$ and .

We next define a binary relationship of inclusion between IF-sets.

Definition 2.53 *Let A and B be twoIF-sets. A is contained in B, and it is denoted by $A \quad B$, if*

$$\mu_A(\omega) \leq \mu_B(\omega) \text{ and } \nu_A(\omega) \geq \nu_B(\omega) \text{ for any } \omega \quad \Omega.$$

Example 2.54 *Let us considera possibility space Ω representing aset of three cities: city 1, city 2 and cit y 3. Let P be a polit ician, and let us consider the IF-sets:*

$A=$ *"P is a good politician".*
$B=$ *"P is honest".*
$C=$ *"P is close to the people".*

Since A, B and C areIF-sets, each cityhas a degreeofagreementwith feature A, B and C, and a degreeof disagreement. In Figure 2.5 wecan seethe membership and non-membership functions of these IF-sets.

Now, inorder to compute the intersectionof the IF-sets A and B,

$$A \cap B = \text{"P is a good politician and honest".}$$

we must compute the value of $\mu_{A \cap B}$ and

$$
\begin{aligned}
\mu_{A \cap B}(city\ i) &= \min(\mu_A(city\ i), \mu_B(city\ i)) = \mu_B(city\ i), \quad \text{for } i = 1, 2, 3. \\
\nu_{A \cap B}(city\ i) &= \max(\nu_A(city\ i), \nu_B(city\ i)) = \nu_B(city\ i), \quad \text{for } i = 1, 2, 3.
\end{aligned}
$$

Thus, $A \cap B = B$. It holds since $B \quad A$, in the sense that $\mu_B \leq \mu_A$ and $\nu_B \geq \nu_A$, and its interpretation would be that P is less honest than a good politician.

Now, let us computethe IF-set "P is honest or closeto the people", that is, the IF-set $B \quad C$. We obtain that:

$$\mu_{B \quad C}(city\ i) = \max(\mu_B(city\ i), \mu_C(city\ i)) = \begin{cases} \mu_B(city\ i) & \text{for } i = 1, 3. \\ \mu_C(city\ i) & \text{for } i = 2. \end{cases}$$

$$\nu_{B \cap C}(city\ i) = \min(\nu_B(city\ i), \nu_C(city\ i)) = \begin{cases} \nu_B(city\ i) & \text{for } i = 1, 3. \\ \nu_C(city\ i) & \text{for } i = 2. \end{cases}$$

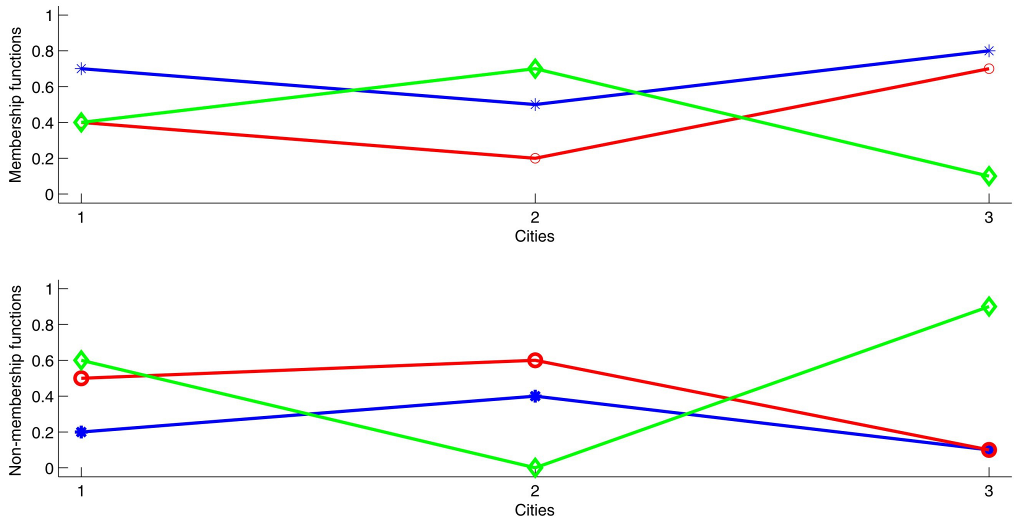


Figure 2.5: Examples of the memb ership and non-memb ership functions of the IF-sets that express the P is a go od p olitician( ), Pishonest(°) and P is close to the p eople (♦).

Then, the IF-set B C *can be expressed in the fol lowing way:*

$$B \quad C = \{(city\ 1, \mu_B(city\ 1), \nu_B(city\ 1)),$$
$$(city\ 2, \mu_C(city\ 2), \nu_C(city\ 2)), (city\ 3, \mu_B(city\ 3), \nu_B(city\ 3))\}.$$

Let us conclude this part by defining the difference op erator b etween IF-sets. According to [27], a difference between fuzzy sets, or *fuzzy difference*, is a map $FS(\Omega) \times FS(\Omega) \to FS(\Omega)$ such that for every pair of fuzzy sets A and B it satisfies the following prop erties:

If A B, then $A - B =$.
If A A, then $A - B$ $A - B$.

Some examples of fuzzy differences are the following:

$$A - B(\omega) = \max\{0, A(\omega) - B(\omega)\},$$
$$A - B(\omega) = \begin{cases} A(\omega) & \text{if } B(\omega) = 0, \\ 0 & \text{otherwise,} \end{cases}$$

for any ω Ω.

Similarly, we can extend the defi nition of difference for IF-sets.

Definition 2.55 *An operator* $-$: $IF\ Ss(\Omega) \times IF\ Ss(\Omega) \rightarrow IF\ Ss(\Omega)$ *isa difference between IF-set s (*IF-difference*, in short) if it satisfies properties D1 and D2.*

D1 If A B, *then* $A - B =$.
D2 If A A , *then* $A - B$ $A - B$.

Any function D satisfying D1 and D2 is a difference op erator. Nevertheless,there are other interesting prop erties that IF-differences may satisfy:

D3 $(A \cap C) - (B \cap C)$ $A - B$.
D4 $(A \quad C) - (B \quad C)$ $A - B$.
D5 $A - B =$ A B.

Letus give an example of IF-difference that alsofulfills D3, D4 andD5.

Example 2.56 *Consider thefunction* $-$: $IF\ Ss(\Omega) \times IF\ Ss(\Omega) \rightarrow IF\ Ss(\Omega)$ *given by:*

$$A - B = \{(\omega, \mu_{A-B}(\omega), \nu_{A-B}(\omega)) | \omega \quad \Omega\},$$

where

$$\mu_{A-B}(\omega) = \max(0, \mu_A(\omega) - \mu_B(\omega));$$

$$\nu_{A-B}(\omega) = \begin{cases} 1 - \mu_{A-B}(\omega) & \text{if } \nu_A(\omega) > \nu_B(\omega); \\ \min(1 + \nu_A(\omega) - \nu_B(\omega), 1 - \mu_{A-B}(\omega)) & \text{if } \nu_A(\omega) \leq \nu_B(\omega). \end{cases}$$

Let us provethat this funct ion satisfies properties D1 and D2, i.e., that it is an IF-difference.

D1: Let us take A B. *Then* $\mu_A \leq \mu_B$ *and* $\nu_A \geq \nu_B$.

$$\mu_{A-B}(\omega) = \max(0, \mu_A(\omega) - \mu_B(\omega)) = 0.$$
$$\nu_{A-B}(\omega) = 1 - \mu_{A-B}(\omega) = 1, \text{ because } \nu_A \geq \nu_B.$$

As aconsequence, $A - B =$.

D2: Consider A A , *that is,* $\mu_A \leq \mu_A$ *and* $\nu_A \geq \nu_A$, *and let us prove that* $A - B$ $A - B$. *Thus, forany* ω *in* Ω *we have that:*

$$\mu_{A-B}(\omega) = \max(0, \mu_A(\omega) - \mu_B(\omega)) \leq \max(0, \mu_A(\omega) - \mu_B(\omega)) = \mu_{A-B}(\omega).$$

$$\nu_{A-B}(\omega) = \begin{cases} 1 - \mu_{A-B}(\omega) & \text{if } \nu_A(\omega) > \nu_B(\omega). \\ \min(1 - \mu_{A-B}(\omega), 1 + \nu_A(\omega) - \nu_B(\omega)) & \text{if } \nu_A(\omega) \leq \nu_B(\omega). \end{cases}$$
$$\leq \begin{cases} 1 - \mu_{A-B}(\omega) & \text{if } \nu_A(\omega) > \nu_B(\omega). \\ \min(1 - \mu_{A-B}(\omega), 1 + \nu_A(\omega) - \nu_B(\omega)) & \text{if } \nu_A(\omega) \leq \nu_B(\omega). \end{cases}$$
$$\leq \nu_{A-B}(\omega).$$

This shows that $-$ *is an IF-difference. Letus see thatit also satisfiesproperties D3, D4 and D5.*

D3: *Let us take into account that the IF-sets* $A \cap C$ *and* $B \cap C$ *are given by:*

$$A \cap C = \{(\omega, \min(\mu_A(\omega), \mu_C(\omega)), \max(\nu_A(\omega), \nu_C(\omega))) \mid \omega \quad \Omega\}$$
$$B \cap C = \{(\omega, \min(\mu_B(\omega), \mu_C(\omega)), \max(\nu_B(\omega), \nu_C(\omega))) \mid \omega \quad \Omega\}.$$

For short, we will denote by D *the IF-set* $D = A \cap C - B \cap C$. *On one hand, we are going to prove that* $\mu_{A-B} \geq \mu_D$:

$$\mu_{A-B}(\omega) = \max(0, \mu_A(\omega) - \mu_B(\omega)).$$
$$\mu_D(\omega) = \max(0, \min(\mu_A(\omega), \mu_C(\omega)) - \min(\mu_B(\omega), \mu_C(\omega))).$$

Applying the first part of Lemma A.1 of Appendix A, we deduce that $\mu_{A-B} \geq \mu_D$.

Now, let us prove that $\nu_{A-B} \leq \nu_D$. *There are two possibilities, either* $\nu_A(\omega) > \nu_B(\omega)$ *or* $\nu_A(\omega) \leq \nu_B(\omega)$. *Assume that* $\nu_A(\omega) > \nu_B(\omega)$. *In such a case,* $\max(\nu_A(\omega), \nu_C(\omega)) \geq \max(\nu_B(\omega), \nu_C(\omega))$ *and* $\nu_{A-B}(\omega) = 1 - \mu_{A-B}(\omega)$, *and consequently:*

$$\nu_D(\omega) = 1 - \mu_D(\omega) \geq 1 - \mu_{A-B}(\omega) = \nu_{A-B}(\omega).$$

Assume now that $\nu_A(\omega) \leq \nu_B(\omega)$. *Then it holds that*

$$\max(\nu_A(\omega), \nu_C(\omega)) \leq \max(\nu_B(\omega), \nu_C(\omega)).$$

By the second part of Lemma A.1 of Appendix A,

$$\nu_B(\omega) - \nu_A(\omega) \geq \max(\nu_B(\omega), \nu_C(\omega)) - \max(\nu_A(\omega), \nu_C(\omega)),$$

whence

$$\begin{aligned}\nu_D(\omega) &= \min(1 + \max(\nu_A(\omega), \nu_C(\omega)) - \max(\nu_B(\omega), \nu_C(\omega)), 1 - \mu_D(\omega)) \\ &\geq \min(1 + \nu_A(\omega) - \nu_B(\omega), 1 - \mu_{A-B}(\omega)) = \nu_{A-B}(\omega).\end{aligned}$$

Thus we conclude that $\nu_{A-B} \leq \nu_D$, *and therefore* $(A \cap C) - (B \cap C) \quad A - B$.

D4: *Consider three IF-sets* A, B *and* C. *The IF-sets* $A \quad C$ *and* $B \quad C$ *are given by:*

$$A \quad C = \max(\mu_A, \mu_C), \min(\nu_A, \nu_C).$$
$$B \quad C = \max(\mu_B, \mu_C), \min(\nu_B, \nu_C).$$

Let us denote by D *the IF-set* $D = (A \quad C) - (B \quad C)$, *and let us prove that* $\mu_{A-B} \geq \mu_D$. *This is equivalent to*

$$\max(0, \mu_A(\omega) - \mu_B(\omega)) \geq \max(0, \max(\mu_A(\omega), \mu_C(\omega)) - \max(\mu_B(\omega), \mu_C(\omega))),$$

for every $\omega \quad \Omega$, *and this inequality holds because of the first part of Lemma A.1 of Appendix A.*

Let us prove that $\nu_D \geq \nu_{A-B}$. *To see this, consider the two possible cases:* $\nu_A(\omega) > \nu_B(\omega)$ *and* $\nu_A(\omega) \leq \nu_B(\omega)$. *Assume that* $\nu_A(\omega) > \nu_B(\omega)$, *which means that* $\nu_{A-B}(\omega) =$

$1 - \mu_{A-B}(\omega)$. *Now,* $\nu_A(\omega) > \nu_B(\omega)$ *implies that* $\min(\nu_A(\omega), \nu_C(\omega)) \geq \min(\nu_B(\omega), \nu_C(\omega))$ *and therefore:*

$$\nu_D(\omega) = 1 - \mu_D(\omega) \geq 1 - \mu_{A-B}(\omega) = \nu_{A-B}(\omega).$$

Assume now that $\nu_A(\omega) \leq \nu_B(\omega)$, *whence*

$$\min(\nu_A(\omega), \nu_C(\omega)) \leq \min(\nu_B(\omega), \nu_C(\omega)).$$

Applying thesecond partof Lemma A.1 of AppendixA, we knowthat

$$\nu_B(\omega) - \nu_A(\omega) \geq \min(\nu_B(\omega), \nu_C(\omega)) - \min(\nu_A(\omega), \nu_C(\omega)).$$

Then, we deduce that:

$$\begin{aligned} \nu_D(\omega) &= \min(1 + \min(\nu_A(\omega), \nu_C(\omega)) - \min(\nu_B(\omega), \nu_C(\omega)), 1 - \mu_D(\omega)) \\ &\geq \min(1 + \nu_A(\omega) - \nu_B(\omega), 1 - \mu_A(\omega)) = \nu_{A-B}(\omega). \end{aligned}$$

Thus, $\nu_D \geq \nu_{A-B}$, *and therefore* $(A \quad C) - (B \quad C) \quad A - B$.

D5: Let us consider A *and* B *such that* $A - B = \quad$. *Then,* $\mu_{A-B}(\omega) = 0$ *and* $\nu_{A-B}(\omega) = 1$ *for every* $\omega \quad \Omega$, *whence*

$$0 = \mu_{A-B}(\omega) = \max(0, \mu_A(\omega) - \mu_B(\omega)) \quad \mu_A(\omega) \leq \mu_B(\omega).$$

$$1 = \nu_{A-B}(\omega) = \begin{cases} 1 & \text{if } \nu_A(\omega) > \nu_B(\omega). \\ 1 + \nu_A(\omega) - \nu_B(\omega) & \text{if } \nu_A(\omega) \leq \nu_B(\omega). \end{cases}$$

Therefore, $\mu_A(\omega) \leq \mu_B(\omega)$ *and* $\nu_A(\omega) \geq \nu_B(\omega)$, *and as a consequence* $A \quad B$.

3 Comparison of alternatives underun-certainty

This memory is devoted to the comparison of alternatives under some lack of information. If this lack of information is given by uncertainty ab out the consequences of the alternatives, these are usually mo delled by means of random variables. Thus, sto chastic orders emerge as an essential to ol, since they allow the comparison of random quantities. As we mentioned in the previous chapter, one of the most imp ortant sto chastic orders in the literature is that of sto chastic dominance, in any of its degrees. Sto chastic dominance has b ee n widely investigated (see [98, 108, 109, 110, 173], among others) and it has been applied in many different areas ([11, 77, 95, 109, 171, 180]). However, the other sto chastic order we have intro duced, statistic al preference, has been studied in ([14, 15, 16, 49, 54, 55, 56, 57, 58]) but not as widely as sto chastic dominance. For this reason, the first ste p of this chapter is to make a thorough study of statistical preference. First of all, we investigate its basic prop erties as a sto chastic order, and then we study its relationship with sto chastic dominance. In this sense, we shall firstly lo ok for conditions that guarantee that first degree sto chastic dominance implies statistical preference. Then, we shall show that in general there is not an implication relationship between statistical preference and the n-th degree sto chastic dominance. We als o provide several examples of the b ehaviour of statistical preference, and also sto chastic dominance, in some of the most usual distributions, like for instance Bernoulli, exp onential or, of course, the normal distribution.

Both sto chastic dominance and statistical preference are stochastic orders that were intro duced for the pairwise comparison of random variables. In fact, statistical preference presents a disadvantage that is its lack of transitivity, as was pointed out by several authors ([14, 15, 16, 49, 54, 56, 58, 121, 122]). To illustrate this fact, we give an example. Then, in order to have an sto chastic order that allows for the simultaneous comparison of more than two random variables, we present a generalisation of statistic al preference, and study some of its prop erties. In particular, we shall see its connections with the metho ds established for pairwise comparisons.

It is obvious that sto chastic orders are powerful to ols for comparing uncertain quantities. For this reason, and in order to illus trate our results, we conclude the chapter by mentioning two possible applications. On the onehand, we investigate both sto chastic dominance and statistical preference as metho ds for the comparison of fitness values ([180, 183]), and on the other hand we illustrate the usefuln ess of b oth statistical preference and its generalisation for the comparison of more than two random variablesin multicriteria decision making problems with linguistic lab els ([123]).

3.1 Properties of the statistical preference

This section is devoted to the study of the main prop erties of statistical preference. In particular, weshall try tofindacharacterization of this notion: ona firststep, a similar one to that of sto chastic dominance presented in Theorem 2.10;afterwards, we explain that statistical preference seems to be closer to another lo cation parameter, the median.

3.1.1 Basic prop erties and intuitive interpretation of the statistical preference

We start this subsection with some basic prop erties ab out the b ehaviour of the statistical preference.

Lemma 3.1 *Let X and Y be tworandom variables. Thenit holds that*

$$X \quad_{SP} Y \qquad P(X<Y) \le \frac{1}{2}.$$

In part icu lar, the converse implication holds for random variables with $P(X=Y)=0$.

Pro of It holdsthat $Q(X,Y) = P(X>Y) + \frac{1}{2}P(X=Y) \ge \frac{1}{2}$. Then:

$$P(X<Y) = 1 - P(X>Y) - P(X=Y) \le \frac{1}{2} - \frac{1}{2}P(X=Y) \le \frac{1}{2}.$$

If $P(X=Y)=0$,then:

$$Q(X,Y) = P(X>Y) = 1 - P(Y>X) \ge \frac{1}{2},$$

since we assume $P(X<Y) \le \frac{1}{2}$. Thus, $X \quad_{SP} Y$. ∎

Remark 3.2 *Note thatthe converse implicationof theprevious resultdoes nothold in general. Asa counterexample, itisenoughtoconsidertheindependentrandomvariables* *defined by:*

X	0	2
P_X	0.8	0.2

Y	0	1
P_Y	0.7	0.3

On the one hand, it holds that:

$$P(X < Y) = P(X = 0, Y = 1) = P(X = 0)P(Y = 1) = 0.8 \quad 0.3 = 0.24 < \frac{1}{2},$$

and also

$$P(X = Y) = P(X = 0, Y = 0) = P(X = 0)P(Y = 0) = 0.7 \quad 0.8 = 0.56.$$

However, $P(X > Y) = P(X = 2) = 0.2$. *Thus:*

$$Q(X, Y) = P(X > Y) + \frac{1}{2}P(X = Y) = 0.2 + \frac{1}{2} \quad 0.56 = 0.48 < \frac{1}{2}.$$

Now we present a result that shows how tran slations and dilations or contractions affect to the b ehavi ou r of statistical preference for real-valued random variables.

Prop osition 3.3 *Let X, Y and Z be three real-valued random variables defined on the same probability space and let $\lambda = 0$ and μ be two real numbers. It holdsthat*

1. $X \;_{SP}\; Y \qquad X + Z \;_{SP}\; Y + Z$.
2. $\lambda X \;_{SP}\; \mu Y \qquad \begin{array}{ll} X \;_{SP}\; \frac{\mu}{\lambda} Y & \text{if } \lambda > 0. \\ \frac{\mu}{\lambda} Y \;_{SP}\; X & \text{if } \lambda < 0. \end{array}$

Pro of

1. Itholds that

$$\begin{aligned} Q(X, Y) &= P(X > Y) + \frac{1}{2}P(X = Y) \\ &= P(X + Z > Y + Z) + \frac{1}{2}P(X + Z = Y + Z) = Q(X + Z, Y + Z). \end{aligned}$$

Then, $Q(X, Y) \geq \frac{1}{2}$ ifand onlyif $Q(X + Z, Y + Z) \geq \frac{1}{2}$.

2. Let us develop the exp re ssion of $Q(\lambda X, \mu Y)$:

$$Q(\lambda X, \mu Y) = \begin{array}{ll} P X > \frac{\mu}{\lambda} Y + P X = \frac{\mu}{\lambda} Y = Q\, X, \frac{\mu}{\lambda} Y & \text{if } \lambda > 0. \\ P X < \frac{\mu}{\lambda} Y + P X = \frac{\mu}{\lambda} Y = Q\, \frac{\mu}{\lambda} Y, X & \text{if } \lambda < 0. \end{array}$$

Then, the result direct follows from the expres sion of $Q(\lambda X, \mu Y)$. ∎

Some new equivalences can be deduced from the previous ones.

Corollary 3.4 *Let X and Y bea pair ofreal-valuedrandomvariables, λ and μ two real numbers and α aconstant. Then it holds that*

1. $\lambda X \ {}_{SP}\ \mu \quad \begin{cases} X \ {}_{SP}\ \frac{\mu}{\lambda}, & \text{if } \lambda > 0, \\ \frac{\mu}{\lambda} \ {}_{SP}\ X, & \text{if } \lambda < 0, \\ 0 \geq \mu, & \text{if } \lambda = 0. \end{cases}$
2. $X \ {}_{SP}\ Y \quad 1 - Y \ {}_{SP}\ 1 - X.$
3. $X \ {}_{SP}\ Y \quad X - Y \ {}_{SP}\ 0.$
4. $X + Y \ {}_{SP}\ Y \quad X \ {}_{SP}\ 0.$
5. $X \ {}_{SP}\ X + \alpha \quad \alpha \leq 0.$
6. $X \ {}_{SP}\ \alpha X \quad \begin{matrix} 0 \ {}_{SP}\ X, & \text{if } \alpha > 1, \\ X \ {}_{SP}\ 0, & \text{if } \alpha < 1. \end{matrix}$

Pro of In point1, the case of $\lambda > 0$ and $\lambda < 0$ directly follow from item 2 of the previous prop osition. If $\lambda = 0$, applying Remark2.19, the comparison of degenerate random variables is equivalent to the comparison of real numb ers, and then, it is obvious that $\lambda X \ {}_{SP}\ \mu \quad 0 \geq \mu$.

Point 2 follows from the previous prop osition: $X \ {}_{SP}\ Y$ if andonly if $X - 1 \ {}_{SP}\ Y - 1$, and from the second item this is equivalent to $1 - Y \ {}_{SP}\ 1 - X$.

Points 3, 4 and 5 are immediate from the firs t p oint of Prop osition 3.3 and Re-mark 2.19 in the case of 3. Considerthe lastone. Applying our previous prop osition,

$$X \ {}_{SP}\ \alpha X \quad (1 - \alpha) X \ {}_{SP}\ 0.$$

By the second item of Prop osition 3.3,

$$(1 - \alpha) X \ {}_{SP}\ 0 \quad \begin{matrix} 0 \ {}_{SP}\ X, & \text{if } \alpha > 1, \\ X \ {}_{SP}\ 0, & \text{if } \alpha < 1. \end{matrix}$$

Let us compare the b ehaviour of statistical preference and sto chastic dominance with resp ect these basic prop erties.Ontheone hand, Prop osition 2.13 assures that $X_1 + \ldots + X_n \ {}_{FSD}\ Y_1 + \ldots + Y_n$ when the variables are indep endent and $X_i \ {}_{FSD}\ Y_i$. First statement of Prop osition 3.3 assures that $X \ {}_{SP}\ Y \quad X + Z \ {}_{SP}\ Y + Z$, and the indep endence condition is not imposed. However, it is not possible to give a result as general as Prop osition 2.13 for statistical preferenceFor instance, cons ider the universe

$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, with a discrete uniform distribution, and the following random variables:

	ω_1	ω_2	ω_3	ω_4
X_1	-2	1	-2	1
X_2	1	-2	-2	1
Y	0	0	0	0
$X_1 + X_2$	-1	-1	-4	2
$Y + Y$	0	0	0	0

It holds that $X_1 \equiv_{SP} Y$ and $X_2 \equiv_{SP} Y$. However, $Q(X_1 + X_2, Y + Y) = \frac{1}{4}$, and therefore $X_1 + X_2 \ {}_{SP}\ Y + Y$.

First item of Corollary 3.4 trivially holds for stochastic dominance. The second item also holds since:

$$F_{1-X}(t) = 1 - P(X < 1 - t) \text{ and } F_{1-Y}(t) = 1 - P(Y < 1 - t),$$

and then $F_{1-Y}(t) \leq F_{1-X}(t)$ if and only if $P(X < 1 - t) \leq P(Y < 1 - t)$. Note that $P(X \leq t) \leq P(Y \leq t)$ for any t if and only if $P(X < t) \leq P(Y < t)$ for any t: on the one hand, assume that $P(X \leq t) \leq P(Y \leq t)$ for any t. Then:

$$P(X < t) = \lim_{n\to\infty} P\left(X \leq t - \frac{1}{n}\right) \leq \lim_{n\to\infty} P\left(Y \leq t - \frac{1}{n}\right) = P(Y < t).$$

On the other hand, if $P(X < t) \leq P(Y < t)$ for any t, it holds that:

$$P(X \leq t) = \lim_{n\to\infty} P\left(X < t + \frac{1}{n}\right) \leq \lim_{n\to\infty} P\left(Y < t + \frac{1}{n}\right) = P(Y \leq t).$$

We conclude that $X \ {}_{FSD}\ Y$ if and only if $1 - Y \ {}_{FSD}\ 1 - X$. However, stochastic dominance does not satisfy the third item of Corollary 3.4. For instance, if X and Y are two independent and equally distributed random variables following a Bernoulli distribution of parameter $\frac{1}{2}$, it holds that:

$X - Y$	-1	0	1
P_{X-Y}	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Then, $X - Y$ is not comparable with the degenerate variable in 0 for first degree stochastic dominance, but $X \ {}_{FSD}\ Y$.

Furthermore, the fourth item of the previous corollary does not hold, either: it suffices to consider the universe $\Omega = \{\omega_1, \omega_2, \omega_3\}$ with discrete uniform distribution, and the random variables defined by:

	ω_1	ω_2	ω_3
X	0	1	2
Y	2	1	0
$X - Y$	-2	0	2

Then, $X \equiv_{FSD} Y$, but $X - Y$ and 0 are not comparable with respect to stochastic dominance. Nevertheless, first degree stochastic dominance does satisfy the fifth and sixth properties of Corollary 3.4.

Remark 3.5 *Using the third item of the previous corollary, we know that $X \ {}_{SP} \ Y$ if and only if $X - Y \ {}_{SP} \ 0$. This allowed Couso and Sánchez [46] to prove a simple characterization of statistical preference for real-valued random variables:*

$$X \ {}_{SP} \ Y \quad X - Y \ {}_{SP} \ 0 \quad E[u(X - Y)] \geq 0 \tag{3.1}$$

for the function $u: \mathbb{R} \to \mathbb{R}$ defined by $u = I_{(0,\infty)} - I_{(-\infty,0)}$.

Theorem 2.10 showed that $X \ {}_{FSD} \ Y$ if and only if the expectation of $u(X)$ is greater than the expectation of $u(Y)$ for any increasing function u. In particular, Proposition 2.12 assures that, when $X \ {}_{FSD} \ Y$ and ϕ is a increasing function, $\phi(X) \ {}_{FSD} \ \phi(Y)$. In the case of statistical preference, we can check that it is invariant by strictly increasing transformations of the random variables as well.

Proposition 3.6 *Let X and Y be two random variables. It holds that:*

$$X \ {}_{SP} \ Y \quad h(X) \ {}_{SP} \ h(Y)$$

for any strictly order preserving function $h: \Omega \to \Omega$.

Proof On the one hand, if $h(X) \ {}_{SP} \ h(Y)$ for any strictly order preserving function h, by considering the identity function we obtain that $X \ {}_{SP} \ Y$.

On the other hand, note that:

$$\{\omega : h(X(\omega)) > h(Y(\omega))\} = \{\omega : X(\omega) > Y(\omega)\},$$

and consequently $P(X > Y) = P(h(X) > h(Y))$. Similarly, $P(X = Y) = P(h(X) = h(Y))$ and $P(Y > X) = P(h(Y) > h(X))$. Then $Q(X, Y) = Q(h(X), h(Y))$. We conclude that $X \ {}_{SP} \ Y \quad h(X) \ {}_{SP} \ h(Y)$. ∎

However, although first degree stochastic dominance is invariant under increasing transformations, for statistical preference the previous result does not hold for order preserving functions that are not strictly order preserving. For instance, consider the following independent random variables:

X	0	2
P_X	$\frac{1}{2}$	$\frac{1}{2}$

Y	1
P_Y	1

Then, the probabilistic relation takes the value $Q(X, Y) = \frac{1}{2}$. Consider the increasing, but not strictly increasing, function $h: \mathbb{R} \to \mathbb{R}$ given by:

$$h(t) = \begin{cases} t & \text{if } t \quad (-\infty, 0] \quad (2, \infty). \\ 2 & \text{otherwise.} \end{cases}$$

Then, $h(X)$ and $h(Y)$ are given by:

$h(X)$	0	2
$P_{h(X)}$	$\frac{1}{2}$	$\frac{1}{2}$

$h(Y)$	2
$P_{h(Y)}$	1

Thus, $Q(h(X), h(Y)) = \frac{1}{4}$, and then the previous result does not hold.

The last basic property we are going to study is if statistical preference is preserved by different kinds of convergence.

Remark 3.7 *Let $\{X_n\}_n$ and $\{Y_n\}_n$ be two sequences of random variables and let X and Y other two random variables, all of them defined on the same probability space. It holds that:*

$$\left.\begin{array}{l} X_n \xrightarrow{L} X \\ Y_n \xrightarrow{L} Y \\ X_n \succeq_{SP} Y_n \quad \forall n \end{array}\right\} \not\Rightarrow X \succeq_{SP} Y,$$

$$\left.\begin{array}{l} X_n \xrightarrow{P} X \\ Y_n \xrightarrow{P} Y \\ X_n \succeq_{SP} Y_n \quad \forall n \end{array}\right\} \not\Rightarrow X \succeq_{SP} Y,$$

$$\left.\begin{array}{l} X_n \xrightarrow{m-p} X \\ Y_n \xrightarrow{m-p} Y \\ X_n \succeq_{SP} Y_n \quad \forall n \end{array}\right\} \not\Rightarrow X \succeq_{SP} Y,$$

$$\left.\begin{array}{l} X_n \xrightarrow{a.s.} X \\ Y_n \xrightarrow{a.s.} Y \\ X_n \succeq_{SP} Y_n \quad \forall n \end{array}\right\} \not\Rightarrow X \succeq_{SP} Y,$$

where $\xrightarrow{L}$, $\xrightarrow{P}$, $\xrightarrow{m-p}$ and $\xrightarrow{a.s.}$ denote the convergence of random variables in distribution, probability, m^{th}-mean and almost sure, respectively.

It suffices to consider the same counterexample for all the cases: consider the universe $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and the probability P such that $P(\{\omega_1\}) = P(\{\omega_3\}) = \frac{2}{5}$ and $P(\{\omega_2\}) = P(\{\omega_4\}) = \frac{1}{10}$. Let X, X_n, Y and Y_n be the random variables defined by:

	ω_1	ω_2	ω_3	ω_4
X, X_n	0	0	1	1
Y	0	1	1	1
Y_n	$\frac{-1}{n}$	1	1	1

Y_n converges to Y almost surely, and consequently also converges in probability and in distribution. Furthermore, it also converges in m^{th} mean, since:

$$E[(|Y_n - Y|)^m] = \frac{2}{5}\left(\frac{1}{n}\right)^m \xrightarrow{n\to\infty} 0.$$

Also, X_n converges to X for the four kinds of convergence. Furthermore, $X_n \succ_{SP} Y_n$ since:

$$Q(X_n, Y_n) = P(X_n > Y_n) + \frac{1}{2} P(X_n = Y_n)$$
$$= P(\{\omega_1\}) + \frac{1}{2} P(\{\omega_3, \omega_4\}) = \frac{2}{5} + \frac{1}{2}\frac{1}{2} = \frac{13}{20} > \frac{1}{2}.$$

However, $X \nsucc_{SP} Y$, since:

$$Q(X, Y) = P(X > Y) + \frac{1}{2} P(X = Y) = \frac{1}{2} P(\{\omega_1, \omega_3, \omega_4\}) = \frac{9}{20} < \frac{1}{2}.$$

Thus, we can see that, although sto chastic dominance is preserved by the four kind of convergence (see Prop. 2.14), statisticalpreference isnot.

Now we shall trytoclarify themeaning of statistical preference by means of a gambling examp le.

Example 3.8 *Suppose wehave tworandom variables X and Y defined overthe same probability space such that $X \succ_{SP} Y$, i.e., such that $Q(X, Y) > \frac{1}{2}$. Consider the fol lowing game:* weobtain a pairof randomvaluesof X and Y simultaneously. Forexample, if X and Y mo del the results of the dice, we would roll them simultaneously; otherwise, they can b e simu late d by a computer. Player 1 bets 1 euro on Y totake a valuegreater than X. Ifthis holds, Player 1 wins 1 euro, he loses 1 euro if the value of X is greater, and he do es not lose anything if the values are equal.

Denote by Z_i therandom variable“rewardof Player 1 in the i-th iteration of the game”. Thenit holds that

$$Z_i = \begin{cases} 1, & \text{if } Y > X \\ 0, & \text{if } Y = X \\ -1, & \text{if } Y < X \end{cases}$$

Then, applying the hypothesis $P(X > Y) + \frac{1}{2} P(X = Y) > \frac{1}{2}$, it holds that

$$P(X > Y) > \frac{1}{2}(1 - P(X = Y)) = \frac{1}{2}(P(X > Y) + P(Y > X))$$
$$P(X > Y) > P(Y > X),$$

or equivalently, $q > p$, if weconsiderthenotation $p = P(X < Y)$ and $q = P(X > Y)$. Thus

$$E(Z_i) = P(Y > X) - P(Y < X) = p - q < 0.$$

$\{Z_1, Z_2, \ldots\}$ is an infinite sequence of independent and identical ly distributed random variables. Applyingthe large lawof big numbers,

$$\overline{Z_n} = \frac{Z_1 + \ldots + Z_n}{n} \xrightarrow{p} p - q,$$

or equ ivalently,

$$\forall \varepsilon > 0, \lim_{n\to\infty} P\left(|\overline{Z_n} - (p-q)| > \varepsilon\right) = 0. \qquad (3.2)$$

Denote the accu mulated reward of Player 1 after n iterations of the game by S_n. It holds that $S_n = Z_1 + \ldots + Z_n$. Then, Player 1 wins the game after n iterations if $S_n > 0$. Then, taking $\varepsilon = q - p$ in Equation (3.2), Player 1 wins the game after n iterations with probability:

$$\begin{aligned} P(S_n > 0) &= P(Z_1 + \ldots + Z_n > 0) = P(\overline{Z_n} > 0) \\ &= P(\overline{Z_n} - (p-q) > q - p) \leq P(|\overline{Z_n} - (p-q)| > q - p) \\ &= P(|\overline{Z_n} - (p-q)| > \varepsilon). \end{aligned}$$

Then it holds that:

$$\lim_{n\to\infty} P(S_n > 0) \leq \lim_{n\to\infty} P(|\overline{Z_n} - (p-q)| > \varepsilon) = 0.$$

We have proven that the probability of the event: "Player 1 wins after n iterations of the game" goes to 0 when n goes to ∞.

An immediate consequence is the next prop osition:

Prop osition 3.9 *Let X and Y be two random variables such that $X \succ_{SP} Y$. Consider the experiment that consists of drawing a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ of X and Y, and let*

$B_n \equiv$ "In the first n iterations, at least half of the times the value obtained by X is greater than or equal to the value obtained by Y".

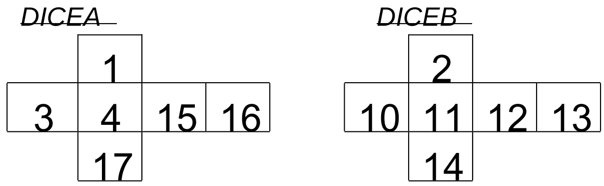
Then,

$$\lim_{n\to\infty} P(B_n) = 1.$$

Then we can say that if we consider the gam e consis ting of obtaining a random value of X and a random value of Y and we rep eat it a large enough numb er of times, if $X \succ_{SP} Y$, we will obtain that more than half of the times the variable X will take a value greater than the value obtained by Y. However, this do es not guarantee that the mean value obtained by the variable X is greater than the mean value obtained by the variable Y.

Let us consider a new example:

Example 3.10 ([57]) *Let us consider the game consisting of rol ling two special dice, denoted A and B, whose results are assumed to be independent. Their faces do not show the classical values but the fol lowing numbers:*



In each iteration, the dice with the greatest number wins.

In this case the probabilistic relation Q *of Equation* (2.7) *takes the value:*

$$
\begin{aligned}
Q(A,B) &= P(A>B) + \frac{1}{2}P(A=B) = P(A>B) = P(A \in \{3,4\}, B \in \{2\}) \\
&\quad + P(A \in \{15,16,17\}, B \in \{2,10,11,12,13,14\}) = \frac{5}{9}.
\end{aligned}
$$

Thus, $A \succ_{SP} B$ *and applying the previous result, if we repeat the game indefinitely, it holds that the probability of winning, betting on* A*, at least half of the times tends to* 1.

However, if we calculate the expected value of every dice, we obtain that

$$
\begin{aligned}
E(A) &= \frac{1}{6}(1+3+4+15+16+17) = \frac{28}{3}, \\
E(B) &= \frac{1}{6}(2+10+11+12+13+14) = \frac{31}{3}.
\end{aligned}
$$

Then, by the criterium of the highest expected reward, dice B *should be preferred. The same applies if we consider the criterion of stochastic dominance. However, if our goal is to win the majority of times then we should choose dice* A*.*

3.1.2 Characterizations of statistical preference

In Subsection 2.1.1 we have seen that stochastic dominance can be characterised by means of the direct comparison of the expectation of adequate transformations of the random variables (see Theorem 2.10). In this subsection we shall give characterisations for statistical preference. For this aim, we distinguish different cases: we start by considering independent random variables, then we consider comonotonic and countermonotonic random variables and we conclude with random variables coupled by means of an Archimedean copula. Finally, we show an alternative characterization of statistical preference in terms of the median. Recall that in the rest of this section, we will consider real-valued random variables.

Independent random variables

We start by considering independent random variables. In order to characterise statistical preference for them, we need this previous result.

Lemma 3.11 *Consider two independent real-valued random variables X and Y whose associated cumulative distribution functions are F_X and F_Y, respectively. Then:*

$$P(X \geq Y) = E[F_Y(X)], \tag{3.3}$$

where $E[h(X)]$ stands for the expectation of the function h with respect to the variable X, this is, $E[h(X)] = \int h(x)dF_X(x)$.

Proof In order to prove this result, we consider [24, Theorem 20.3]: given two random vectors X and Y defined on $\mathbb{R}^j$ and $\mathbb{R}^k$, and whose distribution functions are F_X and F_Y, respectively, it holds that:

$$P((X, Y) \in B) = \int_{\mathbb{R}^j} P((x, Y) \in B)dF_X(x), \quad B \in \mathbb{R}^{j+k}. \tag{3.4}$$

In this case, consider $j = k = 1$ and $B = \{(x, y) : x \geq y\}$. Then:

$$P((X, Y) \in B) = P(X \geq Y) \text{ and}$$
$$P((x, Y) \in B) = P(Y \leq x) = F_Y(x).$$

Then, if we put these values into Equation (3.4), we obtain that $P(X \geq Y) = E[F_Y(X)]$. ■

We can now establish the following result.

Theorem 3.12 *Let X and Y be two independent real-valued random variables defined on the same probability space. Let F_X and F_Y denote their respective cumulative distribution functions. If X' is a random variable identically distributed to X and independent of X and Y, it holds that $X \succeq_{SP} Y$ if and only if:*

$$E[F_Y(X)] - E[F_X(X')] \geq \frac{1}{2}(P(X = Y) - P(X = X')). \tag{3.5}$$

Proof It holds that $X \succeq_{SP} Y$ if and only if $P(Y > X) + \frac{1}{2}P(X = Y) \leq \frac{1}{2}$. On the other hand let us recall (see for example [24, Exercise 21.9(d)]) that $E(F_X(X')) = \frac{1}{2} + \frac{1}{2}P(X = X')$. Then, using also Equation (3.3):

$$\begin{aligned} P(Y > X) &= 1 - P(Y \leq X) = 1 - E[F_Y(X)] \\ &= \frac{1}{2} + E[F_X(X')] - \frac{1}{2}P(X = X') - E[F_Y(X)]. \end{aligned}$$

Whereas, $X \succ_{SP} Y$ if and only if

$$\frac{1}{2} + E(F_X(X)) - \frac{1}{2}P(X = X) - E[F_Y(X)] + \frac{1}{2}P(X = Y) \leq \frac{1}{2},$$

or equivalently,

$$E[F_Y(X)] - E[F_X(X)] \geq \frac{1}{2}(P(X = Y) - P(X = X)).$$ ■

Theorem 3.12 generalises the result established in [54, Equation 12] for continuous and independent random variables. For this particular case, Equation (3.5) can be simplified. The reason is that for continuous and independent random variables X, X and Y the probabilities $P(X = Y)$ and $P(X = X)$ equals zero, and then the second part of Equation (3.5) is simplified.

Corollary 3.13 *Let X and Y be two real-valued independent and continuous random variables with cumulative distribution functions F_X and F_Y, respectively. Then:*

$$X \succeq_{SP} Y \Leftrightarrow E[F_Y(X)] \geq E[F_X(X)].$$

If we are dealing with discrete and independent real-valued random variables, Equation (3.5) can also be re-written. Before showing how, let us give the following lemma:

Lemma 3.14 *Let $\{p_n\}_{n \in \mathbb{N}}$ be a sequence of positive real numbers such that $\sum_n p_n = 1$. Then it holds that:*

$$1 = \sum_n p_n^2 + 2\sum_{n<m} p_n p_m.$$

Proof The result is a direct consequence of:

$$1 = \sum_n p_n = \left(\sum_n p_n\right)^2 = \sum_n p_n^2 + 2\sum_{n<m} p_n p_m.$$ ■

Proposition 3.15 *Let X and Y be two real-valued discrete and independent random variables. If S_X denotes the support of X, then $X \succeq_{SP} Y$ holds if and only if*

$$E[F_Y(X^-) - F_X(X^-)] \geq \frac{1}{2}\sum_{x \in S_X} P(X = x)(P(Y = x) - P(X = x)),$$

where $F_X(t^-)$ and $F_Y(t^-)$ denote the left hand side limit of the cumulative distribution functions F_X and F_Y evaluated in t. That is:

$$F_X(t^-) = P(X < t) \quad \text{and} \quad F_Y(t^-) = P(Y < t).$$

Proof Applying the definition of the probabilistic relation Q:

$$Q(X,Y) = P(X>Y) + \tfrac{1}{2}P(X=Y)$$
$$= \sum_{x\in S_X} P(X=x)P(Y<x) + \frac{1}{2}\sum_{x\in S_X} P(X=x)\,P(Y=x)$$
$$= \sum_{x\in S_X} P(X=x)F_Y(x^-) + \frac{1}{2}\sum_{x\in S_X} P(X=x)\,P(Y=x).$$

Thus, $Q(X,Y) \geq \frac{1}{2}$ if and only if:

$$\sum_{x\in S_X} P(X=x)F_Y(x^-) \geq \frac{1}{2}\left(1 - \sum_{x\in S_X} P(X=x)P(Y=x)\right).$$

Applying Lemma 3.14, the right hand side of the previous inequality becomes:

$$\frac{1}{2}\left(\sum_{x\in S_X} P(X=x)^2 + 2\sum_{x_1,x_2\in S_X, x_1<x_2} P(X=x_1)P(X=x_2)\right.$$
$$\left. - \sum_{x\in S_X} P(X=x)\,P(Y=x)\right) = \frac{1}{2}\left(\sum_{x\in S_X} P(X=x)^2\right.$$
$$\left. + 2\sum_{x\in S_X} P(X=x)F_X(x^-) - \sum_{x\in S_X} P(X=x)P(Y=x)\right)$$
$$= E[F_X(X^-)] + \frac{1}{2}\left(\sum_{x\in S_X} P(X=x)(P(X=x) - P(Y=x))\right).$$

Then, it holds that $Q(X,Y) \geq \frac{1}{2}$ if and only if

$$E[F_Y(X^-) - F_X(X^-)] \geq \frac{1}{2}\sum_{x\in S_X} P(X=x)(P(X=x) - P(Y=x)).$$ ∎

Theorem 3.12 allows to characterise statistical preference between independent random variables. However, we have already said that statistical preference is a method that considers the joint distribution of the random variables. For this reason, we are interested not only in independent random variables but also in dependent ones. Next, we focus on comonotonic and countermonotonic random variables, that correspond to the extreme cases of joint distribution functions according to the Fréchet-Hoeffding bounds given in Equation (2.8).

Continuous comonotonic and countermonotonic random variables

Let us consider two continuous random variables whose cumulative distribution functions are F_X and F_Y, respectively, and f_X and f_Y denote their respective density functions.

First of all, let us consider the case in which X and Y are comonotonic. Then, the joint cumulative distribution function of X and Y is:

$$F_{X,Y}(x,y) = \min(F_X(x), F_Y(y)), \text{ for every } x, y \in \mathbb{R}.$$

The value of the relation $Q(X,Y)$ has already been studied by De Meyer et al.

Proposition 3.16 ([54, Prop.7]) *Let X and Y be two real-valued comonotonic and continuous random variables. The probabilistic relation $Q(X,Y)$ has the following expression:*

$$Q(X,Y) = \int_{x:F_X(x)<F_Y(x)} f_X(x)dx + \frac{1}{2}\int_{x:F_X(x)=F_Y(x)} f_X(x)dx. \tag{3.6}$$

In fact, it holds that:

$$P(X>Y) = \int_{x:F_X(x)<F_Y(x)} f_X(x)dx \text{ and}$$
$$P(X=Y) = \int_{x:F_X(x)=F_Y(x)} f_X(x)dx.$$

Therefore, we obtain that $X \succeq_{SP} Y$ if and only if Equation (3.6) takes a value grater than or equal to $\frac{1}{2}$. However, by Lemma 2.20 we know that $X \succeq_{SP} Y$ if and only if $Q(X,Y) \geq Q(Y,X)$. These are given by:

$$\begin{aligned}
Q(X,Y) &= \int_{x:F_X(x)<F_Y(x)} f_X(x)dx + \frac{1}{2}\int_{x:F_X(x)=F_Y(x)} f_X dx. \\
Q(Y,X) &= \int_{x:F_Y(x)<F_X(x)} f_Y(x)dx + \frac{1}{2}\int_{x:F_Y(x)=F_X(x)} f_Y(x)dx \\
&= 1 - \int_{x:F_X(x)<F_Y(x)} f_Y(x)dx - \frac{1}{2}\int_{x:F_Y(x)=F_X(x)} f_Y(x)dx.
\end{aligned}$$

Hence, we obtain the following:

Corollary 3.17 *Let X and Y be two real-valued comonotonic and continuous random variables, where F_X and F_Y denote their respective cumulative distribution functions and f_X and f_Y denote their respective density functions. Then, $X \succeq_{SP} Y$ if and only if:*

$$\int_{x:F_X(x)<F_Y(x)} (f_X(x)+f_Y(x))dx + \frac{1}{2}\int_{x:F_X(x)=F_Y(x)} (f_X(x)+f_Y(x))dx \geq 1.$$

Assume now that X and Y are continuous and countermonotonic real-valued random variables. In that case, the joint cumulative distribution function is given by:

$$F_{X,Y}(x,y) = \max(F_X(x)+F_Y(y)-1, 0), \text{ for } x, y \in \mathbb{R}.$$

As in the case of comonotonic random variables, De Meyer et al. also found the expression of $Q(X,Y)$.

Proposition 3.18 ([54, Prop.7]) *Let X and Y be two real-valued countermonotonic and continuous random variables. The probabilistic relation $Q(X, Y)$ is given by:*

$$Q(X, Y) = F_Y(u), \tag{3.7}$$

where u is one point that fulfils $F_X(u) + F_Y(u) = 1$.

Therefore, using Equation (3.7) it is possible to state the following proposition.

Proposition 3.19 *Let X and Y be two real-valued countermonotonic and continuous random variables. If F_X and F_Y denote their respective cumulative distribution functions, the following equivalence holds:*

$$X \succeq_{SP} Y \Leftrightarrow F_Y(u) \geq F_X(u),$$

where u is a point such that $F_X(u) + F_Y(u) = 1$.

Proof By definition, $X \succeq_{SP} Y$ if and only if $Q(X, Y) \geq \frac{1}{2}$. However, using Equation (3.7), $Q(X, Y) \geq \frac{1}{2}$ is equivalent to $F_Y(u) \geq \frac{1}{2}$. But, since u satisfies $F_X(u) + F_Y(u) = 1$, $F_Y(u) \geq \frac{1}{2}$ if and only if $F_Y(u) \geq F_X(u)$. ■

Discrete comonotonic and countermonotonic random variables with finite supports

In the previous paragraph we considered continuous comonotonic and countermonotonic random variables, and we characterised statistical preference for them. Now, we also consider real-valued random variables coupled by the minimum or Łukasiewicz operators, but we assume them to be discrete with finite supports. For these variables, De Meyer et al. also found the expression of the probabilistic relation Q.

Proposition 3.20 ([54, Prop. 2]) *Let X and Y be two real-valued comonotonic and discrete random variables with finite supports. Then, their supports, denoted by S_X and S_Y, respectively, can be expressed by:*

$$S_X = \{x_1, \ldots, x_n\} \text{ and } S_Y = \{y_1, \ldots, y_n\}$$

such that $x_1 \leq \ldots \leq x_n$ and $y_1 \leq \ldots \leq y_n$, and such that

$$P(X = x_i) = P(Y = y_i) = P(X = x_i, Y = y_i), \text{ for } i = 1, \ldots, n.$$

Furthermore, the probabilistic relation takes the value:

$$Q(X, Y) = \sum_{i=1}^{n} P(X = x_i)\delta_i^M, \tag{3.8}$$

where

$$\delta_M^i = \begin{cases} 1 & \text{if } x_i > y_i. \\ \frac{1}{2} & \text{if } x_i = y_i. \\ 0 & \text{if } x_i < y_i. \end{cases}$$

The following example illustrates this result.

Example 3.21 ([54, Example 3]) *Consider the comonotonic random variables X and Y defined by:*

X	1	3	4
P_X	0.15	0.4	0.45

Y	2	3	5
P_Y	0.35	0.35	0.3

De Schuymer et al. proved that their supports, S_X and S_Y, respectively, can be expressed by:

$$S_X = \{x_1, x_2, x_3, x_4, x_5\} = \{1, 3, 3, 4, 4\} \text{ and } S_Y = \{y_1, y_2, y_3, y_4, y_5\} = \{2, 2, 3, 3, 5\}$$

and their probabilities can be expressed by:

X	x_1	x_2	x_3	x_4	x_5
P_X	0.15	0.2	0.2	0.15	0.3

Y	y_1	y_2	y_3	y_4	y_5
P_Y	0.15	0.2	0.2	0.15	0.3

Using the notation of the previous result, it holds that:

$$\delta_1^M = 0 \text{ because } x_1 < y_1, \qquad \delta_4^M = 1 \text{ because } x_4 > y_4.$$
$$\delta_2^M = 1 \text{ because } x_2 > y_2, \qquad \delta_5^M = 0 \text{ because } x_5 < y_5.$$
$$\delta_3^M = 0.5 \text{ because } x_3 = y_3.$$

Then:

$$Q(X, Y) = \sum_{i=1}^{5} \delta_i^M P(X = x_i) = P(X = x_2) + \frac{1}{2} P(X = x_3) + P(X = x_4)$$
$$= 0.2 + \frac{1}{2} 0.2 + 0.15 = 0.45.$$

Under the previous conditions, it is possible to define the probability space $(\Omega, \mathcal{P}(\Omega), P_1)$, where $\Omega = \{\omega_1, \ldots, \omega_n\}$ and

$$P_1(\{\omega_i\}) = P(X = x_i), \text{ for any } i = 1, \ldots, n.$$

We can also define the random variables X' and Y' by:

$$X'(\omega_i) = x_i \text{ and } Y'(\omega_i) = y_i \text{ for any } i = 1, \ldots, n.$$

Then, the random variables X' and Y' are equally distributed than X and Y, respectively. This will be a very important fact for results in Section 3.2. Next lemma proves that $Q(X, Y) = Q(X', Y')$.

Lemma 3.22 *Under the previous conditions, it holds that* $Q(X, Y) = Q(X, Y)$.

Proof Let us compute the value of $P(X > Y)$ and $P(X = Y)$:

$$
\begin{aligned}
P_1(X > Y) &= P_1(\{\omega_i : X(\omega_i) = x_i > y_i = Y(\omega_i)\}) \\
&= \sum_{i=1}^{n} P_1(\{\omega_i\}) I_{x_i > y_i} = \sum_{i=1}^{n} P(X = x_i) I_{x_i > y_i}. \\
P_1(X = Y) &= P_1(\{\omega_i : X(\omega_i) = x_i = y_i = Y(\omega_i)\}) \\
&= \sum_{i=1}^{n} P_1(\{\omega_i\}) I_{x_i = y_i} = \sum_{i=1}^{n} P(X = x_i) I_{x_i = y_i}.
\end{aligned}
$$

Then:

$$
\begin{aligned}
Q(X, Y) &= P(X > Y) + \tfrac{1}{2} P(X = Y) \\
&= \sum_{i=1}^{n} P(X = x_i) I_{x_i > y_i} + \frac{1}{2} \sum_{i=1}^{n} P(X = x_i) I_{x_i = y_i} \\
&= \sum_{i=1}^{n} P(X = x_i) \delta_i^M = Q(X, Y).
\end{aligned}
$$

Example 3.23 *Let us continue with Example 3.21. We have two random variables* X *and* Y *and we have seen that their supports can be expressed by* $S_X = \{x_1, \ldots, x_5\} = \{1, 3, 3, 4, 4\}$ *and* $S_Y = \{y_1, \ldots, y_5\} = \{2, 2, 3, 3, 5\}$ *respectively. Their probability distributions are given by:*

X	x_1	x_2	x_3	x_4	x_5
P_X	0.15	0.2	0.2	0.15	0.3

Y	y_1	y_2	y_3	y_4	y_5
P_Y	0.15	0.2	0.2	0.15	0.3

Now, we can define the possibility space $\Omega = \{\omega_1, \ldots, \omega_5\}$, *the probability* P_1 *such that* $P_1(\omega_i) = P(X = x_i)$ *and the random variables* X *and* Y *by:*

$$X(\omega_i) = x_i \text{ and } Y(\omega_i) = y_i \text{ for any } i = 1, \ldots, 5.$$

Now, taking into account that:

$$
\begin{aligned}
x_1 &= 1 < 2 = y_1 & \delta_1^M &= 0 \text{ and } X(\omega_1) < Y(\omega_1), \\
x_2 &= 3 > 2 = y_2 & \delta_2^M &= 1 \text{ and } X(\omega_2) > Y(\omega_2), \\
x_3 &= 3 = y_3 & \delta_3^M &= \tfrac{1}{2} \text{ and } X(\omega_3) = Y(\omega_3), \\
x_4 &= 4 > 3 = y_4 & \delta_4^M &= 1 \text{ and } X(\omega_4) > Y(\omega_4), \\
x_5 &= 4 < 5 = y_5 & \delta_5^M &= 0 \text{ and } X(\omega_5) < Y(\omega_5),
\end{aligned}
$$

it is possible to compute the value of the probabilistic relation $Q(X, Y)$:

$$
\begin{aligned}
Q(X, Y) &= P_1(X > Y) + \frac{1}{2} P_1(X = Y) \\
&= P_1(\{\omega_2, \omega_4\}) + \frac{1}{2} P_1(\{\omega_3\}) = 0.2 + 0.15 + \frac{1}{2} 0.2 = 0.45 < \frac{1}{2},
\end{aligned}
$$

hence Y $_{SP}$ Y. *Furthermore, in Example 3.21 we obtained that* $Q(X, Y) = 0.45$ *and therefore, by the previous lemma, it holds that* $Q(X, Y) = Q(X, Y) = 0.45$

Remark 3.24 *Taking the previous comments into account, we shall assume without loss of generality that any two discrete and comonotonic random variables* X *and* Y *with finite supports are defined in a probability space* $(\Omega, \mathcal{P}(\Omega), P)$ *where* Ω *is finite,* $\Omega = \{\omega_1, \ldots, \omega_n\}$, *and* $X(\omega_i) = x_i$, $Y(\omega_i) = y_i$, *such that* $x_i \leq x_{i+1}$ *and* $y_i \leq y_{i+1}$ *for any* $i = 1, \ldots, n-1$. *Moreover:*

$$P(X = x_i, Y = y_i) = P(X = x_i) = P(Y = y_i) \text{ for } i = 1, \ldots, n.$$

Furthermore, $Q(X, Y)$ *is given by Equation* (3.8).

Next result gives a characterization of statistical preference in terms of the supports of X and Y, and also in terms of the probability measure in the initial space. Its proof is trivial and therefore omitted.

Proposition 3.25 *Consider two real-valued comonotonic and discrete random variables* X *and* Y *with finite supports. According to the previous remark, we can assume them to be defined on* $(\Omega, \mathcal{P}(\Omega), P)$ *where* $\Omega = \{\omega_1, \ldots, \omega_n\}$, *by* $X(\omega_i) = x_i$ *and* $Y(\omega_i) = y_i$, *where* $x_i \leq x_{i+1}$ *and* $y_i \leq y_{i+1}$ *for any* $i = 1, \ldots, n-1$. *Then,* X $_{SP}$ Y *if and only if:*

$$\sum_{i: x_i > y_i} P(X = x_i) \geq \sum_{i: x_i < y_i} P(X = x_i),$$

or equivalently, by Lemma 3.22, if and only if:

$$\sum_{i: x_i > y_i} P(\{\omega_i\}) \geq \sum_{i: x_i < y_i} P(\{\omega_i\}).$$

Now, we focus on countermonotonic random variables. For them, De Meyer et al. proved the following result:

Proposition 3.26 ([54, Prop. 4]) *Let* X *and* Y *be real-valued comonotonic and discrete random variables with finite supports. Then, their supports can be expressed by* $S_X = \{x_1, \ldots, x_n\}$ *and* $S_Y = \{y_1, \ldots, y_n\}$, *respectively, such that* $x_1 \leq \ldots \leq x_n$ *and* $y_1 \leq \ldots \leq y_n$, *and such that:*

$$P(X = x_i) = P(Y = y_{n-i+1}) = P(X = x_i, Y = y_i)$$

for any $i = 1, \ldots, n$. *Under these conditions, the probabilistic relation* $Q(X, Y)$ *takes the value:*

$$Q(X, Y) = \sum_{i=1}^{n} P(X = x_i)\delta_i^{\perp}, \tag{3.9}$$

where

$$\delta_i^L = \begin{cases} 1 & \text{if } x_i > y_{n-i+1}; \\ \frac{1}{2} & \text{if } x_i = y_{n-i+1}; \\ 0 & \text{if } x_i < y_{n-i+1}. \end{cases}$$

To illustrate this result, consider the following example.

Example 3.27 ([54, Example 5]) *Consider the random variables X and Y of Example 3.21, but now assume them to be countermonotonic. Their supports can be expressed by $S_X = \{x_1, x_2, x_3, x_4, x_5\} = \{1, 3, 3, 4, 4\}$ and $S_Y = \{y_1, y_2, y_3, y_4, y_5\} = \{2, 3, 3, 5, 5\}$. Furthermore, the probability distributions of X and Y can be expressed by:*

X	x_1	x_2	x_3	x_4	x_5
P_X	0.15	0.15	0.25	0.1	0.35

Y	y_1	y_2	y_3	y_4	y_5
P_Y	0.35	0.1	0.25	0.15	0.15

Using the notation of the previous result, it holds that:

$\delta_1^L = 0$ *because* $x_1 < y_5$, $\quad$ $\delta_4^L = 1$ *because* $x_4 > y_4$.

$\delta_2^L = 0$ *because* $x_2 < y_4$, $\quad$ $\delta_5^L = 1$ *because* $x_5 > y_5$.

$\delta_3^L = 0.5$ *because* $x_3 = y_3$.

Then:

$$Q(X, Y) = \sum_{i=1}^{5} \delta_i^L P(X = x_i) = \frac{1}{2} P(X = x_3) + P(X = x_4) + P(X = x_5)$$
$$= \frac{1}{2} 0.25 + 0.1 + 0.35 = 0.575.$$

Under the above conditions, and similarly to the case of comonotonic random variables, it is possible to define a probability space $(\Omega, \mathcal{P}(\Omega), P_2)$, where $\Omega = \{\omega_1, \ldots, \omega_n\}$ and the probability is given by:

$$P_2(\{\omega_i\}) = P(X = x_i) \text{ for every } i = 1, \ldots, n.$$

Furthermore, we can also define the random variables X and Y by:

$$X(\omega_i) = x_i \text{ and } Y(\omega_i) = y_{n-i+1} \text{ for any } i = 1, \ldots, n.$$

Note that the variables X and X, and also Y and Y, are equally distributed. Furthermore, next lemma shows that $Q(X, Y) = Q(X, Y)$.

Lemma 3.28 *In the conditions of the previous comments, considering the probability space $(\Omega, \mathcal{P}(\Omega), P_2)$ and the random variables X and Y, it holds that $Q(X, Y) = Q(X, Y)$.*

Proof Let us compute the value of $P_2(X > Y)$ and $P_2(X = Y)$:

$$P_2(X > Y) = P_2(\{\omega_i : X(\omega_i) = x_i > y_{n-i+1} = Y(\omega_i)\})$$
$$= \sum_{i=1}^{n} P_2(\{\omega_i\}) I_{x_i > y_{n-i+1}} = \sum_{i=1}^{n} P(X = x_i) I_{x_i > y_{n-i+1}}.$$
$$P_2(X = Y) = P_2(\{\omega_i : X(\omega_i) = x_i = y_{n-i+1} = Y(\omega_i)\})$$
$$= \sum_{i=1}^{n} P_2(\{\omega_i\}) I_{x_i = y_{n-i+1}} = \sum_{i=1}^{n} P(X = x_i) I_{x_i = y_{n-i+1}}.$$

Then:

$$Q(X, Y) = P_2(X > Y) + \frac{1}{2} P_2(X = Y)$$
$$= \sum_{i=1}^{n} P(X = x_i) I_{x_i > y_{n-i+1}} + \frac{1}{2} \sum_{i=1}^{n} P(X = x_i) I_{x_i = y_{n-i+1}}$$
$$= \sum_{i=1}^{n} P(X = x_i) \delta_i^L = Q(X, Y).$$

Next example helps to understand how to build the probability space and the random variables.

Example 3.29 *Consider again Example 3.27. The supports of the random variables X and Y can be expressed by $S_X = \{x_1, \ldots, x_5\} = \{1, 3, 3, 4, 4\}$ and $S_Y = \{y_1, \ldots, y_5\} = \{2, 3, 3, 5, 5\}$ respectively. Their probability distributions are given by:*

X	x_1	x_2	x_3	x_4	x_5
P_X	0.15	0.15	0.25	0.1	0.35

Y	y_1	y_2	y_3	y_4	y_5
P_Y	0.35	0.1	0.25	0.15	0.15

Now, we can define the possibility space $\Omega = \{\omega_1, \ldots, \omega_5\}$, the probability P satisfying that $P(\{\omega_i\}) = P(X = x_i)$ for $i = 1, \ldots, 5$ and the random variables X and Y by:

$$X(\omega_i) = x_i \text{ and } Y(\omega_i) = y_{6-i} \text{ for any } i = 1, \ldots, 5.$$

Taking into account that:

$$x_1 = 1 < 5 = y_5 \quad \delta_1^L = 0 \text{ and } X(\omega_1) < Y(\omega_1),$$
$$x_2 = 3 < 5 = y_4 \quad \delta_2^L = 0 \text{ and } X(\omega_2) < Y(\omega_2),$$
$$x_3 = 3 = y_3 \quad \delta_3^L = \tfrac{1}{2} \text{ and } X(\omega_3) = Y(\omega_3),$$
$$x_4 = 4 > 3 = y_2 \quad \delta_4^L = 1 \text{ and } X(\omega_4) > Y(\omega_4),$$
$$x_5 = 4 > 2 = y_1 \quad \delta_5^L = 1 \text{ and } X(\omega_5) > Y(\omega_5),$$

it is possible to comput e the value of the probabilistic relation $Q(X, Y)$:

$$Q(X, Y) = P(X > Y) + \frac{1}{2}P(X = Y)$$
$$= \frac{1}{2}P(\{\omega_3\}) + P(\{\omega_4, \omega_5\}) = \frac{1}{2}0.25 + 0.1 + 0.35 = 0.575 \geq \frac{1}{2},$$

whence $Y \ _{SP}\ Y$. *Moreover, from Example 3.27* $Q(X, Y) = 0.575$ *and therefore, as we have seen in the previous lemma,* $Q(X, Y) = Q(X, Y) = 0.575$.

Remark 3.30 *Using the previous result we can assume, without loss of generality, that any two countermonotonic real-valued random variables* X *and* Y *are defined on a probability space* $(\Omega, \mathcal{P}(\Omega), P)$ *where* $\Omega = \{\omega_1, \ldots, \omega_n\}$, *by* $X(\omega_i) = x_i$ *and* $Y(\omega_i) = y_{n-i+1}$ *such that* $x_i \leq x_{i+1}$ *and* $y_i \geq y_{i+1}$ *for* $i = 1, \ldots, n$, *and satisfying that*

$$P(X = x_i, Y = y_i) = P(X = x_i) = P(Y = y_{n-i+1}) \text{ for } i = 1, \ldots, n.$$

Now, assuming the conditions of the previous remark, we prove that there is, at most, one element ω such that $X(\omega_i) = Y(\omega_i)$.

Lemma 3.31 *In the conditions of the previous remark, if there exists* $l > 0$ *such that*

$$X(\omega_k) = \ldots = X(\omega_{k+l}) = Y(\omega_k) = \ldots = Y(\omega_{k+l}),$$
$$\min(|X(\omega_{k-1}) - X(\omega_{k+l+1})|, |Y(\omega_{k-1}) - Y(\omega_{k+l+1})|) > 0,$$

for some k, *then it is possible to define a probability space* $(\Omega, \mathcal{P}(\Omega), P_3)$ *and two random variables* X *and* Y *such that:*

- $Q(X, Y) = Q(X, Y)$.
- *There are not* $\omega, \omega \in \Omega$ *such that*

$$X(\omega) = X(\omega) = Y(\omega) = Y(\omega).$$

- X *and* Y *fol low the same distribution than* X *and* Y, *respectively.*

Pro of Define $\Omega = \{\omega_1, \ldots, \omega_{n-l}\}$ and let P_3 be the probability given by:

$$P_3(\{\omega_i\}) = P(\{\omega_i\}) \text{ for any } i = 1, \ldots, k-1.$$
$$P_3(\{\omega_k\}) = P(\{\omega_k\}) + \ldots + P(\{\omega_{k+l}\}).$$
$$P_3(\{\omega_i\}) = P(\{\omega_{i+1}\}) \text{ for any } i = k+l+1, \ldots, n-1.$$

Consider the random variables X and Y given by:

$$X(\omega_i) = X(\omega_i) \text{ and } Y(\omega_i) = Y(\omega_i) \text{ for any } i = 1, \ldots, k-1.$$
$$X(\omega_k) = X(\omega_k) \text{ and } Y(\omega_k) = Y(\omega_k).$$
$$X(\omega_i) = X(\omega_{i+1}) \text{ and } Y(\omega_i) = Y(\omega_{i+1}) \text{ for any } i = k+l+1, \ldots, n-1.$$

They satisfy that:

$$X(\omega_i) < Y(\omega_i) \text{ for any } i = 1, \ldots, k-1.$$
$$X(\omega_k) = Y(\omega_k).$$
$$X(\omega_i) > Y(\omega_i) \text{ for any } i = k+l+1, \ldots, n-1.$$

Then, since

$$\omega_k \,/\{\omega \quad \Omega : X(\omega) > Y(\omega)\} \text{ and}$$
$$\omega_k, \ldots, \omega_{k+l} \,/\{\omega \quad \Omega : X(\omega) > Y(\omega)\},$$

it holds that:

$$\omega_i \quad \{X > Y\} \qquad \omega_i \quad \{X > Y\}, \text{ for } i = 1, \ldots, k-1.$$

Furthermore, $\omega_{i-1} \,/\{X > Y\}$ and $\omega_i \,/\{X > Y\}$ for $i = 1, \ldots, k-1$. Then, we conclude that:

$$P_3(X > Y) = P_3(\{\omega \quad \Omega : X(\omega) > Y(\omega)\})$$
$$= \quad_{i:X(\omega_i) > Y(\omega_i)} P_3(\{\omega_i\}) = \quad_{i:X(\omega_i) > Y(\omega_i)} P(\{\omega_i\}) = P(X > Y).$$

Furthermore, since $X(\omega_k) = Y(\omega_k)$ and $P_3(\{\omega_k\}) = P(\{\omega_k, \ldots, \omega_{k+l}\})$, it hol ds that:

$$P_3(X = Y) = P_3(\{\omega \quad \Omega : X(\omega) = Y(\omega)\})$$
$$= \quad_{i:X(\omega_i) = Y(\omega_i)} P_3(\{\omega_i\}) = P_3(\{\omega_k\})$$
$$= P(\{\omega_k\}) + \ldots + P(\{\omega_{k+l}\}) = \quad_{i:X(\omega_i) = Y(\omega_i)} P(\{\omega_i\}) = P(X = Y).$$

Then, $Q(X, Y) = Q(X, Y)$.

Moreover, by construction therearenot $\omega, \omega \quad \Omega$, $\omega = \omega$,such that

$$X(\omega) = X(\omega) = Y(\omega) = Y(\omega).$$

Finally, it is obvi ou s that X and X, and also Y and Y, are equally distributed, since they take the samevalues withthe same probabilities. ∎

Remark 3.32 *Takinginto account the previous result and Remark 3.30, we conclude that given twodiscrete countermonotonic random variables X and Y with finit e supports, we canassume, without loss ofgenerality, that their supports are given by $S_X = \{x_1, \ldots, x_n\}$ and $S_Y = \{y_1, \ldots, y_n\}$, where $x_i \leq x_{i+1}$ and $y_i \leq y_{i+1}$ for $i = 1, \ldots, n-1$, and that they are defined in aprobability space $(\Omega, P(\Omega), P)$ where $\Omega = \{\omega_1, \ldots, \omega_n\}$, by $X(\omega_i) = x_i$ and $Y(\omega_i) = y_{n-i+1}$. Furthermore:*

$$P(X = x_i, Y = y_i) = P(X = x_i) = P(Y = y_{n-i+1}) \text{ for any } i = 1, \ldots, n.$$

Under these conditions, $Q(X, Y)$ is givenby Equation (3.9). *Furthermore, using the previous lemma we can also assume that $\max\{|X(\omega_i) - X(\omega_{i+1})|, |Y(\omega_i) - Y(\omega_{i+1})|\} > 0$ for any $i = 1, \ldots, n-1$.*

These results allow us to characterise statistical preference for discrete countermonotonic random variables with finite supports.

Proposition 3.33 *Let X and Y be two real-valued discrete and countermonotonic random variables with finite supports, that can be expressed as in the previous remark. Then, it is possible to characterise $X \succeq_{SP} Y$ in the following way:*

- *If there exists k such that $X(\omega_k) = Y(\omega_k)$, then $X \succeq_{SP} Y$ if and only if:*

$$P(X = x_1) + \ldots + P(X = x_{k-1}) \leq P(X = x_{k+1}) + \ldots + P(X = x_n),$$

  or equivalently, if and only if:

$$P(\{\omega_1\}) + \ldots + P(\{\omega_{k-1}\}) \leq P(\{\omega_{k+1}\}) + \ldots + P(\{\omega_n\}).$$

- *If $X(\omega_i) = Y(\omega_i)$ for any $i = 1, \ldots, n$, denote by $k = \min\{i : X(\omega_i) < Y(\omega_i)\}$. Then $X \succeq_{SP} Y$ if and only if:*

$$P(X = x_1) + \ldots + P(X = x_k) \leq P(X = x_{k+1}) + \ldots + P(X = x_n),$$

  or equivalently, if and only if:

$$P(\{\omega_1\}) + \ldots + P(\{\omega_k\}) \leq P(\{\omega_{k+1}\}) + \ldots + P(\{\omega_n\}).$$

Proof Assume that there is k such that $X(\omega_k) = Y(\omega_k)$. Then, $X(\omega_i) > Y(\omega_i)$ for any $i<k$ and $X(\omega_i) < Y(\omega_i)$ for any $i>k$. Then:

$$Q(X, Y) = P(\{\omega_{k+1}, \ldots, \omega_n\}) + \frac{1}{2}P(\{\omega_k\}) \text{ and}$$
$$Q(Y, X) = P(\{\omega_1, \ldots, \omega_{k-1}\}) + \frac{1}{2}P(\{\omega_k\}).$$

Then, $Q(X, Y) \geq \frac{1}{2}$ if and only if:

$$P(\{\omega_{k+1}, \ldots, \omega_n\}) \geq P(\{\omega_1, \ldots, \omega_{k-1}\}).$$

Furthermore, the previous expression is equivalent to:

$$P(X = x_{k+1}) + \ldots + P(X = x_n) \geq P(X = x_1) + \ldots + P(X = x_{k-1}).$$

Now, assume that $X(\omega_i) = Y(\omega_i)$ for any $i = 1, \ldots, n$. Then, denote by k the element $k = \max\{i : X(\omega_i) < Y(\omega_i)\}$. Then, $X(\omega_i) > Y(\omega_i)$ for any $i = k+1, \ldots, n$ and $X(\omega_i) < Y(\omega_i)$ for any $i = 1, \ldots, k$. Then:

$$Q(X, Y) = P(\{\omega_{k+1}, \ldots, \omega_n\}) \text{ and } Q(Y, X) = P(\{\omega_1, \ldots, \omega_k\}).$$

Then, $Q(X, Y) \geq \frac{1}{2}$ if and only if:

$$P(\{\omega_{k+1}, \ldots, \omega_n\}) \geq P(\{\omega_1, \ldots, \omega_k\}).$$

This expression is equivalent to:

$$P(X = x_{k+1}) + \ldots + P(X = x_n) \geq P(X = x_1) + \ldots + P(X = x_k).$$ ■

Random variables coupled by a strict Archimedean copula

Consider two continuous real-valued random variables X and Y with cumulative distribution functions F_X and F_Y, respectively. Let us denote their density functions by f_X and f_Y, respectively. We shall assume the existence of a strict Archimedean copula C, generated by the twice differentiable generator ϕ, such that

$$F_{X,Y}(x, y) = \phi^{-1]}(\phi(F_X(x)) + \phi(F_Y(y))), \text{ for every } x, y \in \mathbb{R}.$$

Note that since C is strict, then $\phi(0) = \infty$. In that case, we have already mentioned in Equation (2.10) that the pseudo-inverse becomes the inverse, and then the joint cumulative distribution function is given by:

$$F_{X,Y}(x, y) = \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(y))), \text{ for every } x, y \in \mathbb{R}.$$

Now, we are going to obtain the joint density function for (X, Y). For this aim, we derive $F_{X,Y}$ with respect to x and y:

$$\begin{aligned}
\frac{\partial F_{X,Y}}{\partial x}(x, y) &= \frac{\partial \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(y)))}{\partial x}(x, y) \\
&= \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(y)))\,\phi(F_X(x)) f_X(x). \\
\frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x, y) &= \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(y)))\phi(F_X(x))\phi(F_Y(y)) f_X(x) f_Y(y).
\end{aligned}$$

Then, the function $f_{X,Y}$ defined by:

$$f_{X,Y}(x, y) = \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(y)))\phi(F_X(x))\phi(F_Y(y)) f_X(x) f_Y(y), \tag{3.10}$$

is a density function of (X, Y). Let us check that $f_{X,Y}(x, y) \geq 0$ for every $x, y \in \mathbb{R}$:

- $f_X, f_Y \geq 0$ because they are density functions.
- By Definition 2.26, $-\phi$ is 2-monotone. Then, $(-1)^2(-\phi) = -\phi \geq 0$, that implies $\phi \leq 0$. Then, $\phi(F_X(x))\phi(F_Y(y)) \geq 0$.

- Since $-\phi$ is 2-monotone, $(-1)^3(-\phi) \geq 0$, and then $\phi \geq 0$. Also, it is known that, for a function g, $g^{-1}(x) = g(g^{-1}(x))^{-1}$. Then:

$$\phi^{-1}(x) = \frac{1}{\phi(\phi^{-1}(x))},$$

and since $\phi \leq 0$, it holds that $\phi^{-1}(x) \leq 0$. Then:

$$\phi^{-1}(x) = -\frac{\phi(\phi^{-1}(x))\,\phi^{-1}}{\phi(\phi^{-1}(x))}.$$

The denominator is positive because it is squared. Furthermore, ϕ is positive, but ϕ^{-1} is negative, but when multiplying for (-1) it becomes positive.

Then, f is the product of positive elements, and therefore f is positive. Now, let us see that the area below $f_{X,Y}$ is 1:

$$\begin{aligned}
\int_{\mathbb{R}}\int_{\mathbb{R}} f_{X,Y}(x,y)\,dy\,dx &= \int_{\mathbb{R}} \phi^{-1}\left(\phi(F_X(x)) + \phi(F_Y(y))\right)\Big|_{-\infty}^{\infty} \phi(F_X(x)) f_X(x)\,dx \\
&= \int_{\mathbb{R}} \phi^{-1}\left(\phi(F_X(x))\right)\phi(F_X(x)) f_X(x)\,dx \\
&= \phi^{-1}(\phi(F_X(x)))\Big|_{-\infty}^{\infty} = F_X(x)\Big|_{-\infty}^{\infty} = 1.
\end{aligned}$$

Using the expression of the joint density function in Equation (3.10) we can prove the following characterization of the statistical preference.

Theorem 3.34 *Let X and Y be two real-valued continuous random variables, and let F_X and F_Y denote their respective cumulative distribution functions, and f_X and f_Y are their respective density functions. If they are coupled by a strict Archimedean copula C generated by the twice differentiable function ϕ, then $X \succeq_{SP} Y$ if and only if:*

$$E\left[\phi^{-1}\left(\phi(F_X(X)) + \phi(F_Y(X))\right) - \phi^{-1}\left(2\phi(F_X(X))\right)\phi(F_X(X))\right] \geq 0. \quad (3.11)$$

Proof First of all, note that (X, Y) is a continuous random vector with density function $f_{X,Y}$. Then, $P(X = Y) = 0$, and therefore $Q(X, Y) = P(X > Y)$ and $Q(Y, X) = P(Y > X)$.

Denote by A the set $A = \{(x, y) \mid x > y\}$. Then,

$$P(X > Y) = \int_A f_{X,Y}(x,y)\,dy\,dx.$$

Thus,

$$P(X>Y) = \int\int_A f_{X,Y}(x,y)\,dy\,dx = \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{X,Y}(x,y)\,dy\,dx$$
$$= \int_{-\infty}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(y)))\Big|_{-\infty}^{x} \phi(F_X(x)) f_X(x)\,dx$$
$$= \int_{-\infty}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(x)))\phi(F_X(x)) f_X(x)\,dx.$$

Furthermore, it holds that

$$\int_{-\infty}^{\infty} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x)) f_X(x)\,dx$$
$$= \frac{1}{2}\phi^{-1}(2\phi(F_X(X)))\Big|_{-\infty}^{\infty} = \frac{1}{2}(\phi^{-1}(0) - \phi^{-1}(\infty)) = \frac{1}{2}.$$

Therefore, $Q(X,Y) = P(X>Y) \geq \frac{1}{2}$ if and only if

$$E\left[\phi^{-1}(\phi(F_X(X)) + \phi(F_Y(X)))\phi(F_X(X))\right]$$
$$= \int_{-\infty}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(x)))\phi(F_X(x)) f_X(x)\,dx$$
$$\geq \frac{1}{2} = \int_{-\infty}^{\infty} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x)) f_X(x)\,dx$$
$$= E\left[\phi^{-1}(2\phi(F_X(X)))\phi(F_X(X))\right].$$

Hence, this inequality is equivalent to

$$E\left[\left(\phi^{-1}(\phi(F_X(X)) + \phi(F_Y(X))) - \phi^{-1}(2\phi(F_X(X)))\right)\phi(F_X(X))\right] \geq 0. \quad \blacksquare$$

This result holds in particular when the random variables are independent, that is, when the copula that links the variables is the product. We have seen in Section 2.1.2 that the product is a strict Archimedean copula with generator $\phi(t) = -\log t$. In this case:

$$\phi(t) = \frac{-1}{t},\ \phi^{-1}(t) = e^{-t} \text{ and } \phi^{-1} = -e^{t}.$$

By replacing these values in Equation (3.11), we obtain that:

$$\phi^{-1}(\phi(F_X(X)) + \phi(F_Y(X))) - \phi^{-1}(2\phi(F_X(X)))$$
$$= -\exp\{\log F_X(X) + \log F_Y(X)\} + \exp\{2\log F_X(X)\}$$
$$= F_Y(X)F_X(X) - F_X(X)^2.$$

Then, Equation (3.11) becomes:

$$E\left(F_Y(X)F_X(X) - F_X(X)^2\right)\frac{1}{F_X(X)} = E[F_Y(X) - F_X(X)] \geq 0.$$

Thus, weconclude thatforcontinuousrandomvariables X and Y, $X \succeq_{SP} Y$ if andonly if $E[F_Y(X) - F_X(X)] \geq 0$. This result has already been obtained in Corollary 3.13.

Random variables coupled by a nilp otent Archimedean copula

Let us study now the case where the copula that links the real-valued random variabl es is a nilp otent Archimedean copula generated by a twice differentiable generatorIn such case, as we saw in Equati on (2.9) the joint distribution function of X and Y isgiven by:

$$F_{X,Y}(x, y) = \begin{cases} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(y))) & \text{if } \phi(F_X(x)) + \phi(F_Y(y)) \in [0, \phi(0)). \\ 0 & \text{otherwise.} \end{cases}$$

Recall that thi s function cannot b e derived in the p oints (x, y) such that $\phi(F_X(x)) + \phi(F_Y(y)) = \phi(0)$. However, thevalue of $\frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x, y)$ can be computed for the points (x, y) fulfilling $\phi(F_X(x)) + \phi(F_Y(y)) \in [0, \phi(0))$. In fact, the value of this fun ction is:

$$\frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x, y) = \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(y)))\phi(F_X(x))\phi(F_Y(y))f_X(x)f_Y(y).$$

Inthis way,the function $f_{X,Y}$ defined by:

$$f_{X,Y}(x, y) = \begin{cases} \frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x, y) & \text{if } \phi(F_X(x)) + \phi(F_Y(y)) \in [0, \phi(0)), \\ 0 & \text{otherwise,} \end{cases}$$

is ajoint density function of X and Y: on theonehand, $f_{X,Y}$ is a positive function:

$$\left.\begin{array}{l} f_X, f_Y \geq 0 \\ \phi \leq 0 \quad \phi(F_X(x))\phi(F_Y(y)) \geq 0 \\ \phi^{-1} \geq 0 \end{array}\right\} \Rightarrow f_{X,Y} \geq 0,$$

since it is the pro duct of positive functions. On theotherhand,it holdsthat

$$\int_{\mathbb{R}}\int_{\mathbb{R}} f_{X,Y}(x, y)\,dy\,dx = 1.$$

In order to prove the last equality, we intro duce the following notation:

$$y_x = \inf\{y \mid \phi(F_X(x)) + \phi(F_Y(y)) \in [0, \phi(0))\}, \text{ for every } x \in \mathbb{R}.$$
$$s_X = \inf\{x \mid F_X(x) > 0\}.$$

Therefore,

$$\{(x, y) \mid x > s_X, y > y_x\} = \{(x, y) \mid \phi(F_X(x)) + \phi(F_Y(y)) < \phi(0)\}.$$

This implies that:

$$\begin{aligned}
\int_{\mathbb{R}}\int_{\mathbb{R}} f_{X,Y}(x, y)\,\mathrm{d}y\,\mathrm{d}x
&= \int_{s_X}^{\infty}\int_{y_x}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(y)))\phi(F_X(x))\phi(F_Y(y))f_X(x)f_Y(y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{s_X}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(y)))\Big|_{y_x}^{\infty} \phi(F_X(x))f_X(x)\,\mathrm{d}x \\
&= \int_{s_X}^{\infty} \phi^{-1}(\phi(F_X(x)))\phi(F_X(x))f_X(x)\,\mathrm{d}x \\
&= \phi^{-1}(\phi(F_X(x)))\Big|_{s_X}^{\infty} = F_X(x)\Big|_{s_X}^{\infty} = 1 - F_X(s_X) = 1.
\end{aligned}$$

We conclude that $f_{X,Y}$ is a joint density function of X and Y. Let us introduce the following notation:

$$\bar{x} = \inf\{x \mid y_x < x\}. \tag{3.12}$$

Using the function $f_{X,Y}$ and the previous notation, we can prove the following characterization of the statistical preference for random variables coupled by a nilpotent Archimedean copula.

Theorem 3.35 *Let X and Y be two real-valued continuous random variables coupled by a nilpotent Archimedean copula whose generator ϕ is twice differentiable and ϕ is not the zero function. $X \succeq_{SP} Y$ if and only if*

$$\int_{\bar{x}}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(x)))\phi(F_X(x))f_X(x)\,\mathrm{d}x \geq \int_{\bar{x}}^{\infty} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X(x)\,\mathrm{d}x.$$

Proof From Theorem 3.34, (X, Y) is a continuous random vector with joint density functions $f_{X,Y}$. Then, $P(X = Y) = 0$, and consequently $Q(X, Y) = P(X > Y)$ and $Q(Y, X) = P(Y > X)$.

Letus compute the value of $Q(X, Y) = P(X > Y)$.

$$P(X > Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{X,Y}(x, y)\,dy\,dx$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{x} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(y)))\phi(F_X(x))\phi(F_Y(y))f_X(x)f_Y(y)\,dy\,dx$$
$$= \int_{-\infty}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(y)))\Big|_{y_x}^{x} \phi(F_X(x))f_X(x)\,dx$$
$$= \int_{-\infty}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(x)))\,\phi(F_X(x))f_X(x)\,dx.$$

Furthermore, if we denote by x the point

$$x = \inf\{x \mid 2\phi(F_X(x)) \leq \phi(0)\}, \tag{3.13}$$

it holds that:

$$\int_{x}^{\infty} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X(x)\,dx = \frac{1}{2}\phi^{-1}(2\phi(F_X(x)))\Big|_{x}^{\infty} = \frac{1}{2}.$$

For this reason, as $X \succeq_{SP} Y$ ifand onlyif $Q(X, Y) \geq \frac{1}{2}$, then $X \succeq_{SP} Y$ if and onl y if

$$\int_{-\infty}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(x)))\phi(F_X(x))f_X(x)\,dx \geq$$
$$\frac{1}{2} = \int_{x}^{\infty} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X(x)\,dx.$$

Remark 3.36 *Thepreviousremarkdoesnot generalise Proposition 3.19, wherea characterization of statistical preference for continuou s and countermonotonic random variables. The reason is that, althou gh the Łukasiewicz operator is an Archimedean copula, its generator is* $\phi(t) = 1 - t$, *and* $\phi(t) = 0$. *Hence, thiscopula does not satisfy the restriction of the previous theorem, which therefore it is not applicable.*

Characterization of the statistical preference by means of the median

In this section we shall investigate the relationship b etween statistical preferen ce and the well-know notion of median ofa randomvariable. F irs t of all let us show an example to clarify the con nection.

Example 3.37 *Consider againthe random variables ofExample 2.3. Itiseasy tocheck that* $Q(X, Y) = 0.6$ *and therefore* $X \succeq_{SP} Y$. *The intuition here isthat in order toobtain* $Q(X, Y) = 0.6$ c *must be a value greater than* 0 *and smal ler than* 1*; however, the exact value of* $c \in (0, 0.6)$ *is not relevant at al l.*

Thus, in the discrete case, statistical preference orders the values of the support of X and Y, and once they are ordered, the exact value of each point does not matter only its relative position and its probability are important. This idea is similar to that used in the definition of the median.

The first approach to connect statistical preference and the median is to compare the medians of the variables X and Y. Recall that a point t is a median of the random variable X if:

$$P(X \geq t) \geq 0.5 \text{ and } P(X \leq t) \geq 0.5, \tag{3.14}$$

and we denote by $\mathrm{Me}(X)$ the set of medians of the random variable X.

Following the previous example, we conjecture that if the median of X is greater than the median of Y then X should be statistically preferred to Y, and the converse implication should also hold. However, this property does not hold in general.

Remark 3.38 *Let X and Y be two real-valued random variables defined on the same probability space. Then there is not a general relationship between X $_{SP}$ Y and the following statements:*

1. $\mathrm{me}(X) \geq \mathrm{me}(Y)$, *for all* $\mathrm{me}(X)$ $\mathrm{Me}(X)$ *and* $\mathrm{me}(Y)$ $\mathrm{Me}(Y)$.
2. $\mathrm{me}(X) \leq \mathrm{me}(Y)$, *for all* $\mathrm{me}(X)$ $\mathrm{Me}(X)$ *and* $\mathrm{me}(Y)$ $\mathrm{Me}(Y)$.

It is enough to consider the independent random variables X and Y defined in Table 3.1.

X	-2	0	2
P_X	0.4	0.2	0.4

Y	-3	1
P_Y	0.4	0.6

Table 3.1: Definition of random variables X and Y.

Both X and Y have only one median, and they equal to: $\mathrm{me}(X) = 0 < \mathrm{me}(Y) = 1$, *but X $_{SP}$ Y because* $Q(X, Y) = 0.64$.

Since both statistical preference and the comparison of medians are complete relations, the same counterexample allows to show that $\mathrm{me}(X) \geq \mathrm{me}(Y)$ *does not guarantee that X $_{SP}$ Y. Notice that* $\mathrm{me}(Y) \geq \mathrm{me}(X)$. *However,* $Q(Y, X) = 0.36$, *so that Y $_{SP}$ X.*

In order to prove that X $_{SP}$ Y and $\mathrm{me}(X) \leq \mathrm{me}(Y)$ *are not related in general, it is enough to define X as the constant random variable on* 1 *and Y as the constant random variable on* 0. *In this case it is obvious that X and Y have only one median and* $\mathrm{me}(X) > \mathrm{me}(Y)$ *and* $Q(X, Y) = 1$.

We see thus th at statistical preference cannot b e reduced to the comparison of the medians of X, Y . Intere stingly, there is a connection between statistical preference and the median of $X - Y$, aswe shallprove inTheorem3.40. Letus presenta preliminary result.

Prop osition 3.39 *Let X and Y be tworeal-valued randomvariables defined on the same probability space. Then*

$$X \quad _{SP}\ Y \qquad F_{X-Y}(0) \leq F_{Y-X}(0),$$

where F_{X-Y} (respectively, F_{Y-X}) denotes the cumulative distribution function of the random variable $X - Y$ (respectively, $Y - X$).

Pro of By Lemma2.20, $X \quad _{SP}\ Y$ ifandonly if $P(X > Y) \geq P(Y > X)$,but:

$$P(X - Y > 0) \geq P(Y - X > 0) \qquad 1 - F_{X-Y}(0) \geq 1 - F_{Y-X}(0) \qquad F_{X-Y}(0) \leq F_{Y-X}(0).$$

Then, $X \quad _{SP}\ Y$ and $F_{X-Y}(0) \leq F_{Y-X}(0)$ are equivalent. ∎

Therefore, in order to che ck statistical preference it suffices to evaluate the cumulative distribution functions of $X - Y$ and $Y - X$ on 0. Inparticular, if $P(X = Y) = 0$, itsuffices to evaluateoneof thecumulative distributionfunctions, F_{X-Y} on 0, since in this case,

$$Q(X, Y) = 1 - F_{X-Y}(0)$$

and $X \quad _{SP}\ Y$ ifand onlyif $F_{X-Y}(0) \leq \frac{1}{2}$. This equivalence holds in particu lar when the random variablesform a continuous random vector.

We next prove the connection b etwe en statistical preferen ce and the median of $X - Y$.

Theorem 3.40 *Let X and Y betwo real-valuedrandom variablesdefinedonthe same probability space.*

1. $\sup \mathrm{Me}(X - Y) > 0 \qquad X \quad _{SP}\ Y \qquad \sup \mathrm{Me}(X - Y) \geq 0.$
2. $X \quad _{SP}\ Y \qquad \mathrm{Me}(X - Y) \quad [0, \infty).$
3. *The converse implication does not hold, alt hough*

$$\inf \mathrm{Me}(X - Y) > 0 \qquad X \quad _{SP}\ Y.$$

4. *If $P(X = Y) = 0$,then*

$$X \quad _{SP}\ Y \qquad \inf \mathrm{Me}(X - Y) > 0.$$

But even when $P(X = Y) = 0$, $0 \quad \mathrm{Me}(X - Y)$ is not equivalent to $Q(X, Y) = \frac{1}{2}$.

Proof

1. Assume that $\sup \mathrm{Me}(X-Y)>0$. Then, there is a median $me(X-Y)>0$. It holds that:

$$\begin{array}{l} P(X>Y) \geq P(X-Y \geq \mathrm{me}(X-Y)) \geq \frac{1}{2} \\ P(X<Y) \leq P(X-Y<\mathrm{me}(X-Y)) \leq \frac{1}{2} \end{array} \qquad Q(X,Y) \geq Q(Y,X),$$

   and then $X \;_{SP}\; Y$. Assume that $X \;_{SP}\; Y$. Then $P(X \geq Y) \geq P(X \leq Y)$. This implies that $P(X-Y \geq 0) \geq Q(X,Y) \geq \frac{1}{2}$, and therefore there exists a median $me(X-Y) \geq 0$, and therefore $\sup \mathrm{Me}(X-Y) \geq me(X-Y) \geq 0$.

2. By definition, $X \;_{SP}\; Y$ if $Q(X,Y)>\frac{1}{2}$.

   Now, assume $\mathrm{me}(X-Y)<0$ for a median of $X-Y$, then:

$$\frac{1}{2} \geq P((X-Y)>\mathrm{me}(X-Y)) \geq P((X-Y) \geq 0) \geq P(X>Y)+\frac{1}{2}P(X=Y).$$

   A contradiction arises because $Q(X,Y)>\frac{1}{2}$.

3. We first prove the implication. Suppose that $\mathrm{me}(X-Y)>0$ for any $\mathrm{me}(X-Y)$ $\mathrm{Me}(X-Y)$. In such a case:

$$\frac{1}{2} \geq P((X-Y)<\mathrm{me}(X-Y)) \geq P(X-Y \leq 0)=1-P(X>Y).$$

   Hence, $P(X>Y) \geq \frac{1}{2}$ and then $X \;_{SP}\; Y$. Now, assume that $Q(X,Y)=\frac{1}{2}$. In that case, $P(X \geq Y)=P(Y \geq X) \geq \frac{1}{2}$, and then:

$$P(X-Y \geq 0)=P(Y-X \geq 0) \geq \frac{1}{2},$$

   whence $0 \quad \mathrm{Me}(X-Y)$, that contradicts the initial hypothesis.

   Next, we give an example where $X-Y$ has only one median and equals 0, and $Q(X,Y)<\frac{1}{2}$. It is enough to consider the random variables X and Y whose joint mass function is defined on Table 3.2.

X/Y	0	1	2
0	0.1	0	0.4
1	0	0.4	0
2	0	0	0.1

Table 3.2: Definition of random variables X and Y.

   For these variables it holds that $\mathrm{Me}(X-Y)=\{0\}$ but $Y \;_{SP}\; X$, since

$$Q(X,Y)=\frac{1}{2}P((X,Y)=(0,0),(1,1),(2,2))=\frac{1}{2}0.6=0.3<\frac{1}{2}.$$

4. Assume that $P(X = Y) = 0$ and let us prove the equivalence. On the one hand, assume that $X \succ_{SP} Y$. By the second item of this Theorem, we know that every median of $X - Y$ is positive. Assume now that 0 is a median of $X - Y$. Then:

$$\frac{1}{2} \geq P(X - Y > 0) = P(X > Y) = Q(X, Y).$$

Then, $Q(X, Y) \leq \frac{1}{2}$, a contradiction. Assume that, although 0 is not a median of $X - Y$, it is the infimum of the medians. In such a case, there is a point $t > 0$ such that any point in $(0, t]$ is a median of $X - Y$. Then, for any $0 < \varepsilon < t$ it holds that:

$$P(X - Y \geq \varepsilon) \geq \frac{1}{2} \text{ and } P(X - Y \leq \varepsilon) \geq \frac{1}{2}.$$

Then, $P(X - Y \geq 0) \geq P(X - Y \geq \varepsilon) \geq \frac{1}{2}$ and:

$$P(X - Y \leq 0) = F_{X-Y}(0) = \lim_{\varepsilon \to 0} F_{X-Y}(\varepsilon) = \lim_{\varepsilon \to 0} P(X - Y \leq \varepsilon) \geq \frac{1}{2}.$$

This means that 0 is also a median, and we have already seen that this is not possible. We conclude that $\inf \mathrm{Me}(X - Y) > 0$.

On the other hand, we have seen in the third item that when $\inf \mathrm{Me}(X - Y) > 0$, $X \succ_{SP} Y$.

Finally, let us see that if 0 is a median of $X - Y$, even when $P(X = Y) = 0$, this is not equivalent to $Q(X, Y) = \frac{1}{2}$. Consider $\Omega = \{\omega_1, \omega_2\}$, the probability measure given by $P(\{\omega_i\}) = \frac{1}{2}$ for $i = 1, 2$, and the random variables X and Y such that $X(\omega_1) = X(\omega_2) = 0$, $Y(\omega_1) = -1$ and $Y(\omega_2) = 1$. Then, -1 is the only median of $X - Y$, and also -1 is the only median of $Y - X$, but $Q(X, Y) = \frac{1}{2}$ and then $X \equiv_{SP} Y$. On the other hand, consider the space $\Omega = \{\omega_1, \omega_2\}$, $P(\{\omega_1\}) = \frac{3}{4}$ and the random variables defined by:

	ω_1	ω_2
X	0	1
Y	0	0
$X - Y$	0	1

Then, 0 is a median of $X - Y$; however, $Q(X, Y) = \frac{5}{8}$. ∎

This theorem establishes a relationship between statistical preference and the median of the difference of the random variables. The particular case in which $P(X = Y) = 0$ is very useful because in that case statistical preference is characterised by the median. Next, we are going to consider two random variables X and Y, and we are going to show how to modify the variables with the aim of avoiding the case $P(X = Y) > 0$.

Lemma 3.41 *Let X, Y be two real-valued discrete random variables, without points of accumulation on their supports, defined on the same probability space such that* $P(X =$

$Y) > 0$. *Assume thattheir supports* S_X *and* S_Y *can beexpressedby* $S_X = \{x_n\}_n$ *and* $S_Y = \{y_m\}_m$ *such that* $x_n \leq x_{n+1}$ *and* $y_m \leq y_{m+1}$ *for any n,m . In thiscaseit ispossible to build another random variable* X *fulfil ling:*

1. $Q(X, Y) = Q(X, Y)$ *and*

2. $P(X = Y \mid X = x) = 0 \qquad P(X = x) = P(X = x)$.

Pro of We shallusethe followingnotation:

$$P(X = x_n, Y = y_m) = p_{n,m} \text{ for any } n, m.$$

Since $P(X = Y) > 0$, there exists x_n S_X and y_m S_Y such that $x_n = y_m$ and $p_{n,m} > 0$. Then, forany (x_n, y_m) in this situationwe consider $x_n^{(1)}, x_n^{(2)}$ such that:

$$\max\{x_{n-1}, y_{m-1}\} < x_n^{(1)} < x_n = y_m < x_n^{(2)} < \min\{x_{n+1}, y_{m+1}\},$$

where x_{n-1} and x_{n+1} (resp ectively, y_{m-1}, y_{m+1}) denote the preceding and subsequent points of x_n in S_X (resp ectively,of y_m in S_Y), existing b ecaus e since b oth S_X and S_Y have no accumulation points. Let us use the follow ing notation:

$$S_X^a = \{x_n \quad S_X : P(X = x_n, Y = x_n) = 0\}.$$
$$S_X^b = \{x_n \quad S_X : P(X = x_n, Y = x_n) > 0\}.$$

Then, $S_X = S_X^a \quad S_X^b$. Wedefine therandomvariable X whose supp ort is given by:

$$S_X = \{x_n \quad S_X^a\} \quad \{x_n^{(1)}, x_n^{(2)} : x_n \quad S_X^b\}.$$

The joint probability of X and Y is give n by:

$$P(X = x_n, Y = y_m) = p_{n,m} \quad \text{if } x_n \quad S_X^a.$$
$$P(X = x_n^{(1)}, Y = y_m) = P(X = x_n^{(2)}, Y = y_m) = \frac{1}{2} p_{n,m} \quad \text{if } x_n \quad S_X^b.$$

By definition, $P(X = Y) = 0$. Then:

$$
\begin{aligned}
Q(X, Y) = P(X > Y) &= \sum_{x \in S_X} P(X > Y \mid X = x) \\
&= \sum_{x_n \in S_X^a} P(X > Y \mid X = x_n) + \sum_{x_n \in S_X^b} P(X > Y \mid X = x_n^{(1)}) \\
&+ P(X > Y \mid X = x_n^{(2)}) \\
&= \sum_{x_n \in S_X^a} P(X > Y \mid X = x_n) + \sum_{x_n \in S_X^b} \frac{1}{2} P(X > Y \mid X = x_n) \\
&+ \frac{1}{2} \sum_{x_n \in S_X^b} P(X > Y \mid X = x_n) + P(X = x_n, Y = x_n) \\
&= \sum_{x_n \in S_X^a} P(X > Y \mid X = x_n) + \sum_{x_n \in S_X^b} P(X > Y \mid X = x_n) \\
&+ \frac{1}{2} P(X = x_n, Y = x_n) \\
&= \sum_{x_n \in S_X} P(X > Y \mid X = x_n) + \frac{1}{2} \sum_{x_n \in S_X^a} P(X = x_n, Y = x_n) \\
&= P(X > Y) + \frac{1}{2} P(X = Y) = Q(X, Y).
\end{aligned}
$$

This lemma allows us to establish the following theorem.

Theorem 3.42 *Let X and Y be two real-valued discrete random variables on the same probability space, whose supports have no accumulation points and such that $P(X = Y) > 0$. Then $X \succeq_{SP} Y$ if and only if it is possible to find a random variable X in the conditions of Lemma 3.41 such that* $\inf \mathrm{Me}(X - Y) > 0$.

Proof Applying the previous lemma it is possible to build another random variable X such that $Q(X, Y) = Q(X, Y)$, $P(X = Y) = 0$, and if $P(X = Y \mid X = x) = 0$, then $P(X = x) = P(X = x)$.

Therefore, as $P(X = Y) = 0$, by Theorem 3.40 it holds that $X \succeq_{SP} Y$ if and only if $\inf \mathrm{Me}(X - Y) \geq 0$. But since $Q(X, Y) = Q(X, Y)$, it holds that $X \succeq_{SP} Y$ if and only if $\inf \mathrm{Me}(X - Y) \geq 0$. ∎

3.2 Relationship between stochastic dominance and statistical preference

In this section we shall study the relationships between first degree stochastic dominance and statistical preference for real-valued random variables.

We recall once more that sto chastic dominance only uses the marginal distributions of the variables compared. Aswe have seeninSubsection 2.1.2, everyjoint cumulative distribution function is the copula of the marginalcumulative distribution functions. For this reason, as we have already done in the previous subsection, we fo cus on different situations: indep endent, comonotonic and countermonotonic random variables, and random variables coupled by an Archimedean copu la.

Before starting with the main res ults, wearegoing toshow thatingeneral, first degree sto chastic dominance do es not imply statistical preference.

Example 3.43 *Considertherandom variables X and Y whose joint mass probability function is given by:*

$X \setminus Y$	0	1	2
0	0.2	0.15	0
1	0	0.2	0.15
2	0.2	0	0.1

Then, the marginal cumulative distribu tion functions of X and Y are defined by:

	$t<0$	t $[0,1)$	t $[1,2)$	$t \geq 2$
$F_X(t)$	0	0.35	0.7	1
$F_Y(t)$	0	0.4	0.75	1

It fol lows that $X \ {}_{FSD}\ Y$ since $F_X \leq F_Y$. However, $X \ {}_{SP}\ Y$ since:

$$
\begin{aligned}
Q(X,Y) &= P(X>Y)+\frac{1}{2}P(X=Y) \\
&= P(X=2,Y=0)+\frac{1}{2}\ P(X=0,Y=0)+P(X=1,Y=1) \\
&+ P(X=2,Y=2) \quad = 0.2+\frac{1}{2}(0.2+0.2+0.1)=0.45.
\end{aligned}
$$

Thus, $X \ {}_{FSD}\ Y$ does notimply $X \ {}_{SP}\ Y$.

Furthermore, since $X \ {}_{FSD}\ Y$ implies $X \ {}_{nSD}\ Y$ for any $n \geq 2$, the pre vious example also shows that $X \ {}_{nSD}\ Y$ do es not imply $X \ {}_{SP}\ Y$ for any $n \geq 2$.

In the following subsections, we will find sufficient conditions for the implication $X \ {}_{FSD} \quad X \ {}_{SP}\ Y$.

3.2.1 Independent random variables

We start by proving that first degree sto chastic dominance implies statistical preference for indep endent random variables. Forthis aim, takeintoaccount that, when $X \ {}_{FSD}\ Y$, Theorem (2.10) assures that $E[u(X)] \geq E[u(Y)]$ for any increasing function u. In particul ar, if we consider $u=F_Y$, which is an incre asing function, it holdsthat $E[F_Y(X)] \geq E[F_Y(Y)]$. This will b e an interesting fact in order to prove the next res ult.

Theorem 3.44 *Let X and Y be two real-valued independent random variables. Then $X \succeq_{FSD} Y$ implies $X \succeq_{SP} Y$.*

Proof Using Lemma 2.20, it suffices to prove that

$$P(X \geq Y) \geq P(Y \geq X).$$

Since X and Y are independent, by Lemma 3.11 it is equivalent to prove that:

$$E[F_Y(X)] \geq E[F_X(Y)].$$

Moreover, since $X \succeq_{FSD} Y$, $F_X \leq F_Y$, and therefore $E[F_X(Y)] \leq E[F_Y(Y)]$. Thus, it suffices to prove that

$$E[F_Y(X)] \geq E[F_Y(Y)],$$

and this inequality holds because $X \succeq_{FSD} Y$ and then $E[u(X)] \geq E[u(Y)]$ for every increasing function u. ∎

With a similar proof it is possible to establish that the implication holds even when one of the variables strictly dominates the other one. Let us introduce a preliminary lemma.

Lemma 3.45 *Let X and Y be two independent real-valued random variables such that $X \succeq_{FSD} Y$. Then, if $P(Y = t) = 0$ for any t such that $F_X(t) < F_Y(t)$, there exists an interval $[a, b]$ such that $P(Y \in [a, b]) > 0$ and $F_X(t) < F_Y(t)$ for any $t \in [a, b]$*

Proof Let t_0 be a point such that $F_X(t_0) < F_Y(t_0)$. Since both F_X and F_Y are right-continuous,

$$\lim_{\varepsilon \to 0} F_Y(t_0 + \varepsilon) = F_Y(t_0) > F_X(t_0) = \lim_{\varepsilon \to 0} F_X(t_0 + \varepsilon).$$

Then, there is $\varepsilon > 0$ such that:

$$F_X(t_0 + \varepsilon) \leq F_X(t_0) + \frac{F_Y(t_0) - F_X(t_0)}{2} < F_Y(t_0).$$

Considering $\delta = \frac{F_Y(t_0) - F_X(t_0)}{2} > 0$, then $F_Y(t) - F_X(t) \geq \delta > 0$ for any $t \in [t_0, t_0 + \varepsilon]$. We have thus proven that there exists an interval $[a, b]$ such that $F_Y(t) - F_X(t) \geq \delta > 0$ for $t \in [a, b]$. Now, without loss of generality, we can assume that $F_Y(a - \varepsilon) < F_Y(a)$ for any $\varepsilon > 0$ (otherwise, since F_Y is right-continuous, take the point $a = \inf(t : F_Y(t) = F_Y(a))$). Then, since $P(Y = a) = 0$, there exists $\varepsilon > 0$ such that $F_Y(t) - F_X(t) \geq \delta > 0$ for any $t \in [a - \varepsilon, b]$. Furthermore:

$$P(Y \in [a - \varepsilon, b]) \geq P(Y \in [a - \varepsilon, a]) \geq P(Y \in (a - \varepsilon, a]) = F_Y(a) - F_Y(a - \varepsilon) > 0,$$

and this completes the proof. ∎

The following result had already been established in [14, Proposition 15.3.5]. However, the authors only gave a proof for continuous random variables. Here, we provide a proof for any pair of random variables X and Y.

Proposition 3.46 *Let X and Y be two real-valued independent random variables. Then, $X \succeq_{FSD} Y$ implies $X \succeq_{SP} Y$.*

Proof We have proven in Theorem 3.44 that $E[F_Y(X)] \geq E[F_Y(Y)]$ when $X \succeq_{FSD} Y$. Then, if we prove that $E[F_X(Y)] < E[F_Y(Y)]$ we would obtain that:

$$P(X \geq Y) = E[F_Y(X)] \geq E[F_Y(Y)] > E[F_X(Y)] = P(Y \geq X),$$

and consequently $X \succeq_{SP} Y$.

Let us prove that if $X \succeq_{FSD} Y$, then $E[F_X(Y)] < E[F_Y(Y)]$. By hypothesis, $F_X(t) \leq F_Y(t)$ for every t, and there is t_0 such that $F_X(t_0) < F_Y(t_0)$.

Let us consider two cases. On the one hand, let us assume that $P(Y = t_0) > 0$. In such a case:

$$\begin{aligned} E[F_X(Y)] &= \int F_X \, dF_Y = \int_{\mathbb{R}\setminus\{t_0\}} F_X \, dF_Y + \int_{\{t_0\}} F_X \, dF_Y \\ &\leq \int_{\mathbb{R}\setminus\{t_0\}} F_Y \, dF_Y + P(Y = t_0) F_X(t_0) \end{aligned}$$

On the other hand, assume that there is not t_0 satisfying both $F_X(t_0) < F_Y(t_0)$ and $P(Y = t_0) > 0$. Applying the previous lemma, there is an interval $[a, b]$ such that $F_Y(t) - F_X(t) \geq \delta > 0$ and $P(Y \in [a, b]) > 0$. Then:

$$\begin{aligned} E[F_X(Y)] &= \int F_X \, dF_Y = \int_{\mathbb{R}\setminus[a,a+\varepsilon]} F_X \, dF_Y + \int_{[a,a+\varepsilon]} F_X \, dF_Y \\ &\leq \int_{\mathbb{R}\setminus[a,a+\varepsilon]} F_Y \, dF_Y + \int_{[a,a+\varepsilon]} (F_Y - \delta) \, dF_Y \\ &= \int F_Y \, dF_Y - \delta P(Y \in [a, a+\varepsilon]) < E[F_Y(Y)]. \end{aligned}$$

A similar result was proven in [210] for probability dominance (see Remark 2.22); nevertheless, that result was only valid for continuous random variables.

3.2.2 Continuous comonotonic and countermonotonic random variables

Let X and Y be two random variables with respective cumulative distribution functions F_X and F_Y, and respective density functions f_X and f_Y.

First of all, let us study the relationship between first degree sto chastic dominance and statistical preference for comonotonic rand om variables.

Theorem 3.47 *Let X and Y betwo real-valuedcomonotonicand continuousrandom variables. If $X \succeq_{FSD} Y$, then $X \succeq_{SP} Y$.*

Pro of InCorollary3.17wehave seenthat $X \succeq_{SP} Y$ ifandonly if

$$\int_{x:F_X(x)<F_Y(x)} (f_X(x)+f_Y(x))\mathrm{d}x + \frac{1}{2}\int_{x:F_X(x)=F_Y(x)} (f_X(x)+f_Y(x))\mathrm{d}x \geq 1.$$

However, by hyp othesis $F_X(x) \leq F_Y(x)$ for any $x \in \mathbb{R}$. Then, $\{x : F_X(x) \leq F_Y(x)\} = \mathbb{R}$, and therefore:

$$\begin{aligned}
&\int_{x:F_X(x)<F_Y(x)} (f_X(x)+f_Y(x))\mathrm{d}x + \frac{1}{2}\int_{x:F_X(x)=F_Y(x)} (f_X(x)+f_Y(x))\mathrm{d}x \\
&= \int_{x:F_X(x)\leq F_Y(x)} (f_X(x)+f_Y(x))\mathrm{d}x - \frac{1}{2}\int_{x:F_X(x)=F_Y(x)} (f_X(x)+f_Y(x))\mathrm{d}x \\
&= \int_{\mathbb{R}} (f_X(x)+f_Y(x))\mathrm{d}x - \frac{1}{2}\int_{x:F_X(x)=F_Y(x)} (f_X(x)+f_Y(x))\mathrm{d}x \\
&\geq \int_{\mathbb{R}} (f_X(x)+f_Y(x))\mathrm{d}x - 1 = 2 - 1 = 1.
\end{aligned}$$

Thus, X isstatistically preferredto Y. ∎

Prop osition 3.46 assures that for independent random variables,whenfirst degree sto chastic dominance holds in the strict sense, statistical preference is also strictAs we shall see,this also holdsfor continuous and comonotonicreal-valuedrandom variables. in order to establish this, we give firs t the following lemma.

Lemma 3.48 *Let X and Y betwocontinuous real-valuedrandom variables. Then, if $X \succ_{FSD} Y$, there exists an interval $[a, b]$ such that $F_X(t) < F_Y(t)$ for any $t \in [a, b]$ and $P(X \in [a, b]) > 0$.*

Pro of From the proof of Lemma 3.45 we deduce that there is an interval $[a, b]$ such that $F_Y(t) - F_X(t) \geq \delta > 0$ for any $t \in [a, b]$. Since F_X iscontinuous, thereis $\varepsilon > 0$ such that $F_X(a-\varepsilon) < F_X(a)$ and $F_Y(t) - F_X(t) \geq \frac{\delta}{2} > 0$ for any $t \in [a-\varepsilon, b]$. Then:

$$P(X \in [a-\varepsilon, b]) \geq P(X \in [a-\varepsilon, a]) \geq F_X(a) - F_X(a-\varepsilon) > 0.$$ ∎

Prop osition 3.49 *Let X and Y be two real-valued comonotonic and continuous random variables. If $X \succ_{FSD} Y$, then $X \succ_{SP} Y$.*

Proof On the one hand, since X $_{FSD}$ Y, then X $_{FSD}$ Y, and consequently X $_{SP}$ Y. According to the previous lemma, there is an interval $[a, b]$ such that $F_Y(t) - F_X(t) \geq \delta > 0$ for any t $[a, b]$ and $P(X$ $[a, b]) > 0$. By Lemma 2.20, X $_{SP}$ Y is equivalent to $P(X > Y) > P(Y > X)$, and from Proposition 3.16 this is equivalent to:

$$\ _{x:F_X(x)<F_Y(x)} f_X(x)\mathrm{d}x > \ _{x:F_Y(x)<F_X(x)} f_Y(x)\mathrm{d}x.$$

Now, take into account that the second part of the previous equation equals 0, since $\{x : F_Y(x) < F_X(x)\} =$. In addition:

$$\ _{x:F_X(x)<F_Y(x)} f_X(x)\mathrm{d}x \geq \ _{[a,b]} f_X(x)\mathrm{d}x = P(X \ [a, b]) > 0.$$

Thus, we conclude that X $_{SP}$ Y. ∎

When the random variables are countermonotonic, the relationship between the (non-strict) first degree stochastic dominance and the (non-strict) statistical preference also holds.

Theorem 3.50 *Let X and Y be two real-valued countermonotonic and continuous random variables. If X $_{FSD}$ Y, then X $_{SP}$ Y.*

Proof In Proposition 3.19 we have seen that X $_{SP}$ Y if and only if $F_Y(u) \geq F_X(u)$, where u is one point such that $F_Y(u) + F_X(u) = 1$. However, since X $_{FSD}$ Y, it holds that $F_X(x) \leq F_Y(x)$ for every x $\mathbb{R}$. In particular, it also holds that $F_X(u) \leq F_Y(u)$. ∎

Although it seems intuitive that the same relationship holds with respect to the strict preferences, this is not the case for countermonotonic continuous random variables. To see this, it suffices to consider the countermonotonic random variables X and Y whose cumulative distribution functions of X and Y are defined by:

$$F_X(t) = \begin{cases} 0 & \text{if } t < 0. \\ t & \text{if } t \ [0, 1]. \\ 1 & \text{if } t > 1. \end{cases} \tag{3.15}$$

$$F_Y(t) = \begin{cases} 0 & \text{if } t < -0.1. \\ \frac{1}{2}t + 0.05 & \text{if } t \ [-0.1, 0.1). \\ t & \text{if } t \ [0.1, 1]. \\ 1 & \text{if } t > 1. \end{cases} \tag{3.16}$$

Since $F_X(t) = F_Y(t)$ for any t / $(-0.1, 0.1)$ and $F_X(t) < F_Y(t)$ for t $(-0.1, 0.1)$, it holds that X $_{FSD}$ Y, but $X \equiv_{SP} Y$, since $F_X(u) + F_Y(u) = 1$ for $u = \frac{1}{2}$ and:

$$Q(X, Y) = F_Y(u) = F_Y(0.5) = \frac{1}{2}.$$

$$Q(Y, X) = F_X(u) = F_X(0.5) = \frac{1}{2}.$$

3.2.3 Discrete comonotonic and countermonotonic random variables with finite supports

Let us now assume that X and Y are discrete real-valued random variables with finite support. Then, when these random variables are comonotonic, we obtain the following result:

Theorem 3.51 *If X and Y are two real-valued comonotonic and discrete random variables with finite supports, then $X \succeq_{FSD} Y \Rightarrow X \succeq_{SP} Y$.*

Proof Using Remark 3.24, we can assume w.l.o.g. that X and Y are defined in $(\Omega, \mathcal{P}(\Omega), P)$, where $\Omega = \{\omega_1, \ldots, \omega_n\}$, by $X(\omega_i) = x_i$ and $Y(\omega_i) = y_i$, where $x_i \leq x_{i+1}$ and $y_i \leq y_{i+1}$ for any $i = 1, \ldots, n-1$, and also:

$$P(X = x_i, Y = y_i) = P(X = x_i) = P(Y = y_i) \text{ for any } i = 1, \ldots, n.$$

Moreover, using Proposition 3.25, $X \succeq_{SP} Y$ if and only if

$$\sum_{i: x_i > y_i} P(X = x_i) \geq \sum_{i: x_i < y_i} P(X = x_i).$$

Let us show that $\{i : x_i < y_i\} = \emptyset$ when $X \succeq_{FSD} Y$. Assume that there exists k such that $X(\omega_k) = x_k < y_k = Y(\omega_k)$. Then:

$$F_X(x_k) = P(X \leq X(x_k)) \geq P(\{\omega_1, \ldots, \omega_k\}).$$
$$F_Y(x_k) = P(Y \leq X(x_k)) \leq P(\{\omega_1, \ldots, \omega_{k-1}\}),$$

where last inequality holds since $\omega_k \notin \{Y \leq X(x_k)\}$ because $Y(\omega_k) > X(\omega_k)$. Now, since $X \succeq_{FSD} Y$, it holds that $F_X(x_k) \leq F_Y(x_k)$:

$$P(\{\omega_1, \ldots, \omega_k\}) \leq F_X(x_k) \leq F_Y(x_k) \leq P(\{\omega_1, \ldots, \omega_{k-1}\}).$$

This implies that $P(\{\omega_k\}) = P(\{X = x_k\}) = 0$, but a contradiction arises since $P(\{\omega_k\}) > 0$. Then, we conclude that $\{i : x_i > y_i\} = \emptyset$, and consequently:

$$\sum_{i: x_i > y_i} P(X = x_i) \geq 0 = \sum_{i: x_i < y_i} P(X = x_i).$$

Thus, $X \succeq_{SP} Y$. ∎

Now, it only remains to see that, as for continuous random variables, strict stochastic dominance implies strict statistical preference.

Proposition 3.52 *Let X and Y be two real-valued discrete and countermonotonic random variables with finite supports. Then, $X \succ_{FSD} Y$ implies $X \succ_{SP} Y$.*

Proof It is obvious that $X \;\, _{FSD}\; Y$ implies $X \;\, _{FSD}\; Y$, and then, applying the previous theorem, $X \;\, _{SP}\; Y$ because $\{i : x_i < y_i\} =$. Then, in order to prove that $X \;\, _{SP}\; Y$ it is enough to see that $\{i : x_i > y_i\} =$, that is, there is some k such that $x_k > y_k$.

Since $X \;\, _{FSD}\; Y$, there is some k such that $F_X(y_k) < F_Y(y_k)$. Assume ex-absurdo that $\{i : x_i > y_i\} =$, so $x_i = y_i$ for any $i = 1, \ldots, n$. Since $x_i = y_i$ and $P(X = x_i) = P(Y = y_i) = P(Y = x_i)$, X and Y are equally distributed, and then $X \equiv_{FSD} Y$, a contradiction. ∎

Finally, let us consider discrete countermonotonic random variables with finite supports, and let us see that, in that case, first degree stochastic dominance also implies statistical preference.

Theorem 3.53 *Let X and Y be two real-valued discrete and countermonotonic random variables with finite supports. Then, $X \;\, _{FSD}\; Y$ implies $X \;\, _{SP}\; Y$.*

Proof From remark 3.32, without loss of generality we can assume that X and Y are defined on $(\Omega, \mathcal{P}(\Omega), P)$, where $\Omega = \{\omega_1, \ldots, \omega_n\}$, by $X(\omega_i) = x_i$ and $Y(\omega_i) = y_{n-i+1}$, where $x_i \le x_{i+1}$ and $y_i \le y_{i+1}$ for any $i = 1, \ldots, n-1$, and also:

$$P(X = x_i, Y = y_i) = P(X = x_i) = P(Y = y_{n-i+1}) \text{ for any } i = 1, \ldots, n.$$

Furthermore, we can also assume that

$$\max(|X(\omega_i) - X(\omega_{i+1})|, |Y(\omega_i) - Y(\omega_{i+1})|) > 0 \quad \text{for any } i = 1, \ldots, n-1;$$

and that there exists, at most, one element k such that $X(\omega_k) = Y(\omega_k)$.

In order to prove that $X \;\, _{FSD}\; Y \quad X \;\, _{SP}\; Y$ we consider two cases:

- Assume $X(\omega_i) = Y(\omega_i)$ for any $i = 1, \ldots, n$ and denote $k = \max\{i : X(\omega_i) < Y(\omega_i)\}$. Then, by Proposition 3.33, $X \;\, _{SP}\; Y$ if and only if:

$$P(\{\omega_1\}) + \ldots + P(\{\omega_k\}) \le P(\{\omega_{k+1}\}) + \ldots + P(\{\omega_n\}).$$

  Since $X \;\, _{FSD}\; Y$, $F_X \le F_Y$. Then, taking $\varepsilon = \frac{Y(\omega_k) - X(\omega_k)}{2} > 0$, it holds that:

$$F_X(X(\omega_k)) = P(X \le X(\omega_k)) \ge P(\{\omega_1, \ldots, \omega_k\}).$$
$$F_Y(X(\omega_k)) \le F_Y(Y(\omega_k) - \varepsilon) = P(Y \le Y(\omega_k) - \varepsilon\}) \le P(\{\omega_{k+1}, \ldots, \omega_n\}).$$

- Assume that there is (an unique) k such that $X(\omega_k) = Y(\omega_k)$. Then:

$$F_X(X(\omega_{k-1})) = P(X \le X(\omega_{k-1})).$$
$$F_Y(X(\omega_{k-1})) = P(Y \le Y(\omega_{k-1})).$$

Since $X(\omega_{k-1}) < Y(\omega_{k-1})$, $\omega_{k-1} / \{Y \leq X(\omega_{k-1})\}$, and thisimpliesthat $\{Y \leq X(\omega_{k-1})\} \ \{\omega_k, \omega_{k+1}, \ldots, \omega_n\}$. Furthermore, $\{X \leq X(\omega_{k-1})\} \ \{\omega_1, \ldots, \omega_{k-1}\}$, and then

$$F_X(X(\omega_{k-1})) \geq P(\{\omega_1\}) + \ldots + P(\{\omega_{k-1}\}).$$

We consider two cases:

- Assume that $Y(\omega_k) = X(\omega_{k-1})$. Then $X(\omega_k) = Y(\omega_k) = X(\omega_{k-1})$, and this implies that $\omega_k \ \{X \leq X(\omega_{k-1})\}$. Then:

$$F_X(X(\omega_{k-1})) \geq P(\{\omega_1\}) + \ldots + P(\{\omega_{k-1}\}) + P(\{\omega_k\}).$$
$$F_Y(Y(\omega_{k-1})) = P(\{\omega_k\}) + P(\{\omega_{k+1}\}) + \ldots + P(\{\omega_n\}).$$

  Using that $X \ _{FSD} \ Y$,

$$P(\{\omega_1\}) + \ldots + P(\{\omega_{k-1}\}) \geq P(\{\omega_{k+1}\}) + \ldots + P(\{\omega_n\}).$$

  Applying Prop osition 3.33, $X \ _{SP} \ Y$.

- Ontheotherhand, if $Y(\omega_k) \leq X(\omega_{k-1})$, then it holds that $\{Y \leq X(\omega_{k-1})\} \ \{\omega_{k+1}, \ldots, \omega_n\}$. Henc e:

$$F_Y(X(\omega_{k-1})) = P(Y \leq X(\omega_{k-1})) \leq P(\{\omega_{k+1}\}) + \ldots + P(\{\omega_n\})$$

  and, since $F_X \leq F_Y$ because $X \ _{FSD} \ Y$, it hold s that:

$$\begin{aligned} P(\{\omega_{k+1}\}) + \ldots + P(\{\omega_n\}) &\geq P(Y \leq X(\omega_{k-1})) = F_Y(X(\omega_{k-1})) \\ &\geq F_X(X(\omega_{k-1})) = P(Y \leq X(\omega_1)) \\ &\geq P(\{\omega_{k+1}\}) + \ldots + P(\{\omega_{k-1}\}). \end{aligned}$$

  By Prop osition 3.33, $X \ _{SP} \ Y$. ∎

Unsurprisingly, strict first degree sto chastic dominance do es not imply strict statistical preference, aswe can seeinthe following example:

Example 3.54 *Consider thecountermonotonic randomvariables X and Y defined by:*

X, Y	0	1	2
P_X	0	0.2	0.8
P_Y	0.1	0.1	0.8

For these variables, $X \ _{FSD} \ Y$. From Remark3.32we canassume that X and Y are defined in the probability space $(\Omega, \mathcal{P}(\Omega), P)$ where $\Omega = \{\omega_1, \ldots, \omega_4\}$, and such that:

$P(\{\omega\})$	0.2	0.6	0.1	0.1
Ω	ω_1	ω_2	ω_3	ω_4
X	1	2	2	2
Y	2	2	1	0

Then, $Q(X, Y) = 0.5$, and we conclude that $X \equiv_{SP} Y$.

3.2.4 Random variables coupled by an Archimedean copula

In this subsection we consider two continuous random variables X and Y, with respective cumulative distribution functions F_X, F_Y and with respective density functions f_X and f_Y. We assume that the random variables are coupled by an Archimedean copula C, generated by the twice differentiable function ϕ.

First of all, we consider the case of a strict Archimedean copula. In that case, we also obtain that first degree stochastic dominance implies that statistical preference.

Theorem 3.55 *Let X and Y be two real-valued continuous random variables coupled by a strict Archimedean copula C generated by the twice differentiable function ϕ. Then, $X \succeq_{FSD} Y$ implies $X \succeq_{SP} Y$.*

Proof From Theorem 3.34, $X \succeq_{SP} Y$ if and only if:

$$E\left[\left(\phi^{-1}(\phi(F_X(X)) + \phi(F_Y(X))) - \phi^{-1}(2\phi(F_X(X)))\right)\phi(F_X(X))\right] \geq 0,$$

or equivalently, if

$$\int_{-\infty}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(x)))\phi(F_X(x))f_X(x)\mathrm{d}x \geq \int_{-\infty}^{\infty} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X(x)\mathrm{d}x.$$

This inequality holds because

$$\begin{aligned}
X \succeq_{FSD} Y \quad & F_X(x) \leq F_Y(x) \\
& \phi(F_X(x)) \geq \phi(F_Y(x)) && (\phi \text{ is decreasing}) \\
& 2\phi(F_X(x)) \geq \phi(F_X(x)) + \phi(F_X(x)) \\
& \phi^{-1}(2\phi(F_X(x))) \geq \\
& \phi^{-1}(\phi(F_X(x)) + \phi(F_X(x))) && (\phi^{-1} \text{ is increasing}) \\
& \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X \leq \\
& \phi^{-1}(\phi(F_X(x)) + \phi(F_X(x)))\phi(F_X(x))f_X(x) && (\phi \leq 0.)
\end{aligned}$$

Therefore, X is statistically preferred to Y. ∎

Remark 3.56 *When applying the previous result to the product copula, we obtain that for continuous and independent random variables, $X \succeq_{FSD} Y$ $X \succeq_{SP} Y$. This is not new for us, since Theorem 3.44 states that this relation holds, not only for continuous, but any kind of independent random variables.*

Let us now investigate if such relationship also holds for the strict preference. For this aim, we consider this preliminary lemma.

Lemma 3.57 *Let X and Y betwo continuousrandomvariablessuch that $X \;_{FSD}\; Y$. Then, thereexists aninterval $[a, b]$ such that $F_X(t) < F_Y(t)$ for any $t \quad [a, b]$ and also $P(X \quad [a, b]) > 0$ and*

$$\phi^{-1}\left(\phi(F_X(t)) + \phi(F_Y(t))\right)\phi(F_X(t)) - \phi^{-1}\left(2\phi(F_X(t))\right)\phi(F_X(t)) \geq \delta > 0$$

for any $t \quad [a, b]$

Pro of We have proven in Lemma 3.48 that there exists an interval $[a, b]$ such that $F_Y(t) - F_X(t) \geq \delta > 0$ for any $t \quad [a, b]$ and $P(X \quad [a, b]) > 0$. Then, there isa subinterval $[a_1, b_1]$ of $[a, b]$ where F_X isstrictly increasing.

Now, following thesame steps thanin Theorem 3.55we obtainthat:

$$\begin{aligned} F_X(t) < F_Y(t) \text{ for any } t \quad [a, b] \\ \phi^{-1}\left(\phi(F_X(t)) + \phi(F_Y(t))\right)\phi(F_X(t)) > \\ \phi^{-1}\left(2\phi(F_X(t))\right)\phi(F_X(t)) \text{ for any } t \quad [a_1, b_1]. \end{aligned}$$

Consider $t \quad [a_1, b_1]$ and let

$$\varepsilon = \phi^{-1}\left(\phi(F_X(t)) + \phi(F_Y(t))\right)\phi(F_X(t)) - \phi^{-1}\left(2\phi(F_X(t))\right)\phi(F_X(t)) > 0.$$

Then,there is a subinterval $[a_2, b_2]$ of $[a_1, b_1]$ such that

$$\phi^{-1}\left(\phi(F_X(t)) + \phi(F_Y(t))\right)\phi(F_X(t)) - \phi^{-1}\left(2\phi(F_X(t))\right)\phi(F_X(t)) \geq \frac{\varepsilon}{2} > 0.$$

Furthermore, since F_X is strictlyincreasing in $[a, b]$ itis alsostrictly increasingin $[a_2, b_2]$, and then $P(X \quad [a_2, b_2]) > 0$. ∎

Prop osition 3.58 *Considertwo real-valued continuous random variables X and Y coupled by a strict Archimedean copula C generated by ϕ. Then, $X \;_{FSD}\; Y$ implies $X \;_{SP}\; Y$.*

Pro of We haveto provethat:

$$\int_{-\infty}^{\infty} \phi^{-1}\left(\phi(F_X(x)) + \phi(F_Y(x))\right)\phi(F_X(x)) f_X(x)\,dx > \int_{-\infty}^{\infty} \phi^{-1}\left(2\phi(F_X(x))\right)\phi(F_X(x)) f_X(x)\,dx.$$

Since X and Y arecontinuous, if $X \;_{FSD}\; Y$, then $X \;_{FSD}\; Y$, and consequently $X \;_{SP}\; Y$ by Theorem 3.55. Taking into account the previous lem ma, there exists an interval $[a, b]$ such that $P(X \quad [a, b]) > 0$ and:

$$\phi^{-1}\left(\phi(F_X(t)) + \phi(F_Y(t))\right)\phi(F_X(t)) - \phi^{-1}\left(2\phi(F_X(t))\right)\phi(F_X(t)) \geq \delta > 0$$

for any t $[a, b]$ Then:

$$\int_{-\infty}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(x)))\phi(F_X(x))f_X(x)\mathrm{d}x$$

$$\geq \int_{\mathbb{R}-[a,b]} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X(x)\mathrm{d}x$$

$$+ \int_{[a,b]} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(x)))\phi(F_X(x))f_X(x)\mathrm{d}x$$

$$> \int_{[a,b]} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X(x)\mathrm{d}x$$

$$+ \int_{[a,b]} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X(x)\mathrm{d}x + \int_{[a,b]} \frac{\varepsilon}{2} f_X(x)\mathrm{d}x$$

$$= \int_{-\infty}^{\infty} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X(x)\mathrm{d}x + \frac{\varepsilon}{2} P(X \quad [a, b])$$

$$> \int_{-\infty}^{\infty} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X(x)\mathrm{d}x.$$

Consequently, $X \ _{SP}\ Y$. ∎

Remark 3.59 *As we have already mentioned, in the particular case where the strict Archimedean copula is the product, the relation $X \ _{FSD}\ Y$ $X \ _{SP}\ Y$ was already studied in Proposition 3.46. Such result states the relation not only for continuous, but for every kind of independent random variables.*

It only remains to study the case of nilpotent copulas. In order to do this, we are going to see the following lemma that assures that, over the assumption of $X \ _{FSD}\ Y$, the points $\bar{x}$ and $\underline{x}$, defined on Equations (3.12) and (3.13), respectively, satisfy $\bar{x} \leq \underline{x}$.

Lemma 3.60 *Let X and Y be two real-valued continuous random variables coupled by a nilpotent Archimedean copula C generated by ϕ. If $X \ _{FSD}\ Y$, then it holds that $\bar{x} \leq \underline{x}$.*

Proof First of all, recall that:

$$\underline{x} = \inf\{x \mid 2\phi(F_X(x)) \leq \phi(0)\},$$
$$\bar{x} = \inf\{x : y_x < x\} \text{ and}$$
$$y_x = \inf\{y \mid \phi(F_X(x)) + \phi(F_Y(y)) \quad [0, \phi(0)]\} \text{ for any } x \quad \mathbb{R}.$$

Assume that $\underline{x} < \bar{x}$. Then there exists a point t such that $\underline{x} < t < \bar{x}$ and $y_t > t$. Moreover, from the hypothesis $X \ _{FSD}\ Y$, it holds that

$$F_X(t) \leq F_Y(t) \quad \phi(F_X(t)) \geq \phi(F_Y(t)) \quad t \quad \mathbb{R}.$$

As $x<t$, we know that $2\phi(F_X(t)) < \phi(0)$. Therefore, wehavethat:

$$\phi(F_X(t)) + \phi(F_Y(t)) \le 2\phi(F_X(t)) < \phi(0).$$

Then,

$$y_t = \inf \{y \mid \phi(F_X(t)) + \phi(F_Y(y)) < \phi(0)\} \le t.$$

Therefore, $y_t > t \ge y_t$, a contradiction. Weconclude that $x \ge \bar{x}$ ∎

Using this lemma we can prove that first degree sto chastic dominance also implies statistical preference for continuous random variables coupled by a nilp otent Archimedean copula.

Theorem 3.61 *If X and Y are two real-valued continuous random variables cou pled by anilpotent Archimedean copula whose generator ϕ is twice differentiable suchthat $\phi = 0$, then X $_{FSD}$ Y X $_{SP}$ Y.*

Pro of From Lemma3.60, $\bar{x} \le x$. Furthermore, $F_X(x) \le F_Y(x)$ for every x R. Then, for every $x \ge x$:

$$
\begin{aligned}
X \ _{FSD}\ Y \quad & F_X(x) \le F_Y(x) \\
& \phi(F_X(x)) \ge \phi(F_Y(x)) && (\phi \text{ is decreasing}) \\
& 2\phi(F_X(x)) \ge \phi(F_X(x)) + \phi(F_X(x)) \\
& \phi^{-1}(2\phi(F_X(x))) \ge \\
& \phi^{-1}(\phi(F_X(x)) + \phi(F_X(x))) && (\phi^{-1} \text{ is increasing}) \\
& \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X \le \\
& \phi^{-1}(\phi(F_X(x)) + \phi(F_X(x)))\phi(F_X(x))f_X(x) && (\phi \le 0.)
\end{aligned}
$$

Therefore:

$$
\begin{aligned}
& \int_{\bar{x}}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(x)))\phi(F_X(x))f_X(x)\,dx \\
\ge & \int_{x}^{\infty} \phi^{-1}(\phi(F_X(x)) + \phi(F_Y(x)))\phi(F_X(x))f_X(x)\,dx \\
\ge & \int_{x}^{\infty} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X(x)\,dx \\
\ge & \int_{x}^{\infty} \phi^{-1}(2\phi(F_X(x)))\phi(F_X(x))f_X(x)\,dx.
\end{aligned}
$$

Applying Theorem 3.35, we dedu ce that X $_{SP}$ Y. ∎

Remark 3.62 *Note that this result is not applicable to the Łukasiewicz copula, since its generator is $\phi_W(t) = 1 - t$, and then $\phi(t) = 0$. However, wehave alreadyseen inTheorem 3.50that firstdegree stochastic dominance implies statistical preference for continuous and countermonotonic random variables.*

As in the countermonotonic case, the relationship between the strict preferences does not hold. To see this, consider two continuous random variables X and Y whose cumulative distribution functions are defined in Equations (3.15) and (3.16). If we consider the generator $\phi(t) = 2(1 - t)$, such that $\phi(0) = 2$, there is not (x, y) in the set:

$$\{(x, y) : \phi(F_X(x)) + \phi(F_Y(y)) \quad [0, \phi(0)]\} \quad F_X(t) = F_Y(t),$$

such that either $x \leq 0.1$ or $y \leq 0.1$. Thus, whenever $f_{X,Y} > 0$, $f_{X,Y}$ is symmetric. Then, if (t, t) satisfies $\phi(F_X(t)) + \phi(F_Y(t)) \quad [0, \phi(0)]$, then $F_X(t) = F_Y(t)$. Consequently:

$$P(X > Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f_{X,Y}(x, y) dy\, dx = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f_{X,Y}(y, x) dy\, dx = P(Y > X).$$

and we conclude X and Y are statistically indifferent.

3.2.5 Other relationships between stochastic dominance and statistical preference

In the previous subsection we have seen several conditions under which $X \ _{FSD}\ Y$ implies $X \ _{SP}\ Y$. Now, we analyze if there are other relationships between first and n-th degree stochastic dominance and statistical preference.

We start by proving that statistical preference does not imply neither first nor n-th degree stochastic dominance for any $n \geq 2$.

Remark 3.63 *There exist random variables X and Y such that:*

1. *$X \ _{SP}\ Y$ but $X \ _{nSD}\ Y$, for every $n \geq 1$.*
2. *$X \ _{nSD}\ Y$ but $X \ _{SP}\ Y$, for every $n \geq 2$.*
3. *$X \ _{FSD}\ Y$ but $X \ _{SP}\ Y$.*
4. *$X \ _{FSD}\ Y$, $X \ _{nSD}\ Y$ for any $n \geq 2$ but $X \ _{FSD}\ Y$.*

In Example 3.43 we gave two random variables such that $Y \ _{SP}\ X$ but $X \ _{FSD}\ Y$. Then, $X \ _{nSD}\ Y$ for any $n \geq 1$ and therefore $Y \ _{nSD}\ X$ for any $n \geq 1$. Thus, this is an example where the first and third items hold.

Consider next random variables X and Y such that X follows a uniform distribution in the interval $(10, 11)$ and Y has the following density function:

$$f_Y(x) = \begin{cases} \frac{1}{25} & \text{if } 0 < x < 10, \\ \frac{3}{5} & \text{if } 11 < x < 12, \\ 0 & \text{otherwise.} \end{cases}$$

For these random variables it holds that:

$$Q(X, Y) = P(X > Y) = P(Y < 10) = \frac{2}{5} < \frac{1}{2},$$

and therefore Y $_{SP}$ X. *However, ontheonehand, itistrivial that neither* Y $_{FSD}$ X *nor* X $_{FSD}$ Y. *Moreover,* X $_{nSD}$ Y *for every* $n \geq 2$*:*

$$G_X^2(t) = \begin{cases} 0 & \text{if } t < 10. \\ \frac{(t-10)^2}{2} & \text{if } t \quad [10, 11). \\ t - 10.5 & \text{if } t \geq 11. \end{cases}$$

$$G_Y^2(t) = \begin{cases} 0 & \text{if } t < 0. \\ \frac{t^2}{50} & \text{if } t \quad [0, 10). \\ \frac{2}{5}t - 2 & \text{if } t \quad [10, 11). \\ \frac{1}{10}(343 - 62t + 3t^2) & \text{if } t \quad [11, 12). \\ t - 8.9 & \text{if } t \geq 12. \end{cases}$$

The graphs of these funct ions can be seen in Figure 3.1.

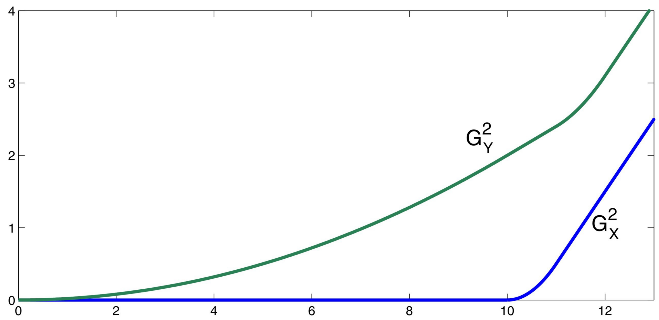


Figure 3.1: Graphics of the fu nctions G_X^2 and G_Y^2.

Then, X $_{SSD}$ Y, *and applying Equation* (2.4), X $_{nSD}$ Y *for every* $n \geq 2$.

We have thus an example where Y $_{SP}$ X *and* X $_{nSD}$ Y *for every* $n \geq 2$.

Letus see bymeans of an examplethat X $_{SP}$ Y *and* X $_{nSD}$ Y *donot guarantee* X $_{FSD}$ Y. *Tosee that, it isenough to consider theindependent randomvariables* X *and* Y *defined by:*

X	1	5
P_X	$\frac{1}{2}$	$\frac{1}{2}$

Y	0	10
P_Y	$\frac{9}{10}$	$\frac{1}{10}$

For these variables it holds that:

$$Q(X, Y) = P(X > Y) + \frac{1}{2}P(X = Y) = P(X > Y) = P(Y = 0) = \frac{9}{10} > \frac{1}{2}.$$

Thus $X \succ_{SP} Y$. *Furthermore, since the cumulative distribution functions are:*

$$F_X(t) = \begin{cases} 0 & \text{if } t < 1, \\ \frac{1}{2} & \text{if } t \in [1, 5), \\ 1 & \text{if } t \geq 5. \end{cases} \qquad F_Y(t) = \begin{cases} 0 & \text{if } t < 0, \\ \frac{9}{10} & \text{if } t \in [0, 10), \\ 1 & \text{if } t \geq 10, \end{cases}$$

the functions G_X^2 *and* G_Y^2 *are:*

$$G_X^2(t) = \begin{cases} 0 & \text{if } t < 1, \\ \frac{1}{2}(t-1) & \text{if } t \in [1, 5), \\ t-3 & \text{if } t \geq 5, \end{cases} \qquad G_Y^2(t) = \begin{cases} 0 & \text{if } t < 0, \\ \frac{9}{10}t & \text{if } t \in [0, 10), \\ t-1 & \text{if } t \geq 10. \end{cases}$$

If we look at their graphical representations in Figure 3.2, we can see that $X \succ_{SSD} Y$. *However,*

$$F_X(5) = 1 > \frac{9}{10} = F_Y(5),$$

whence X *cannot stochastically dominate* Y *by first degree, i.e.,* $X \not\succ_{FSD} Y$.

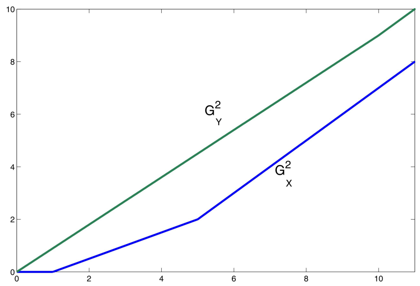


Figure 3.2: Graphics of the functions G_X^2 and G_Y^2.

Our next Theorem summarises the main results of this paragraph.

Theorem 3.64 *Let* X *and* Y *be two random variables.* $X \succ_{FSD} Y$ *implies* $X \succ_{SP} Y$ *under any of the following conditions:*

- X *and* Y *are independent.*

- X and Y are continuous and comonotonic random variables.
- X and Y are continuous and countermonotonic random variables.
- X and Y are discrete and comonotonic random variables with finite supports.
- X and Y are discrete and countermonotonic random variables with finite supports.
- X and Y are continuous random variables coupled by an Archimedean copula.

The relationships between stochastic dominance and statistical preference under the conditions of the previous result are summarised in Figure 3.3.

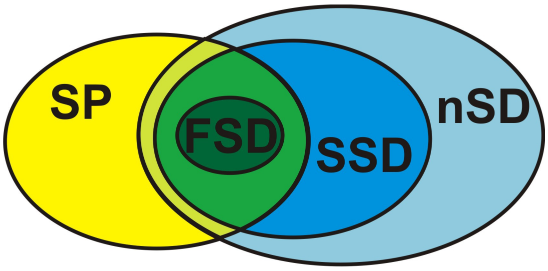


Figure 3.3: General relationship between stochastic dominance and statistical preference.

3.2.6 Examples on the usual distributions

In this subsection we shall study the conditions we must to impose on the parameters of some of the most important parametric distributions in order to obtain statistical preference and stochastic dominance for independent random variables. We shall see that for some of them, stochastic dominance and statistical preference are equivalent. Some results in this sense have already been established in [56].

Discrete distributions under independence: Bernoulli

In the case of discrete distributions, we shall consider the Bernoulli distribution with parameter $p \in (0, 1)$, denoted by $B(p)$, that takes the value 1 with probability p and the value 0 with probability $1 - p$.

Proposition 3.65 *Let X and Y be two independent random variables with distributions $X \equiv B(p_1)$ and $Y \equiv B(p_2)$. Then:*

- $Q(X, Y) = \frac{1}{2}(p_1 - p_2 + 1)$, *and*
- *X is statistically preferred to Y if and only if* $p_1 \geq p_2$.

Proof Let us compute the expression of the probabilistic relation $Q(X, Y)$:

$$\begin{aligned} Q(X, Y) &= P(X > Y) + \tfrac{1}{2} P(X = Y) \\ &= P(X = 1, Y = 0) + \tfrac{1}{2}\; P(X = 0, Y = 0) + P(X = 1, Y = 1) \\ &= p_1(1 - p_2) + \tfrac{1}{2}((1 - p_1)(1 - p_2) + p_1 p_2) = \tfrac{1}{2}(p_1 - p_2 + 1). \end{aligned}$$

Then it holds that:

$$X \quad _{SP}\; Y \qquad Q(X, Y) \geq \frac{1}{2} \qquad \frac{1}{2}(p_1 - p_2 + 1) \geq \frac{1}{2} \qquad p_1 \geq p_2. \qquad \blacksquare$$

Thus, a necessary and sufficient condition for $X \quad _{SP}\; Y$ is that $p_1 \geq p_2$, or equivalently, $E[X] \geq E[Y]$. In fact, it is immediate that this condition is also necessary and sufficient for $X \quad _{FSD}\; Y$. Thus, first degree stochastic dominance is a complete relation for Bernoulli distributions; as a consequence, the same applies to n-th degree stochastic dominance, and therefore they are equivalent methods. This allows us to establish the following corollary.

Corollary 3.66 *Let X and Y be two independent random variables with Bernoulli distribution. Then:*

$$X \quad _{FSD}\; Y \qquad X \quad _{nSD}\; Y \text{ for any } n \geq 2 \qquad X \quad _{SP}\; Y \qquad E[X] \geq E[Y].$$

Continuous distributions under independence

Next, we consider some of the most important families of continuous distributions: exponencial, beta, Pareto and uniform. In addition, due to the importance of the normal distribution, we devote the next paragraph to its study; in that case we shall consider other possibilities in addition to independent random variables.

Remark 3.67 *Although the beta distribution depends on two parameters, $p, q > 0$, in this work we shall consider the particular cases where one of the parameters equals 1, as in [56]. The general case in which both parameters are greater than 1 is much more complex, since the expression of the probabilistic relation is very difficult to obtain.*

Analogously, the Pareto distribution depends on two parameters, a, b, and the density function is given by

$$f(x) = \frac{ab^a}{x^{a+1}}, \qquad x > b.$$

As in [56] we will focus on the case $b = 1$.

Before starting, we recall in Table 3.3the density functionsand theparameters of the distributionswestudy alongthis subsection.

Distribution	Density function	Parameters
Exp onential	$\lambda e^{-\lambda x}, x \in (0, \infty)$	$\lambda > 0$
Uniform	$\frac{1}{b-a}, x \in (a, b)$	$a, b \in \mathbb{R}, a < b$
Pareto	$\lambda x^{-(\lambda+1)}, x \in (1, \infty)$	$\lambda > 0$
Beta	$\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} x^{p-1}(1-x)^{q-1}, x \in (0, 1)$	$p, q > 0$

Table 3.3: Characteristic s of the continuous distributions to b e studied.

Prop osition 3.68 *Let X and Y betwoindependentrandomvariableswith exponential distributions, $X \equiv Exp(\lambda_1)$ and $Y \equiv Exp(\lambda_2)$,respectively. Then:*

- $Q(X, Y) = \frac{\lambda_2}{\lambda_1 + \lambda_2}$ *and*
- *X is statistical ly preferred to Y if andonly if $\lambda_1 \leq \lambda_2$.*

Pro of We firstprove that $Q(X, Y) = \frac{\lambda_2}{\lambda_1 + \lambda_2}$.

$$Q(X, Y) = P(X > Y) = \int_0^\infty \lambda_1 e^{-\lambda_1 x} dx \int_0^x \lambda_2 e^{-\lambda_2 y} dy = \int_0^\infty \lambda_1 e^{-\lambda_1 x}(1 - e^{-\lambda_2 x}) dx$$

$$= \int_0^\infty \lambda_1 e^{-\lambda_1 x} dx - \int_0^\infty \lambda_1 e^{-(\lambda_1 + \lambda_2)x} dx = 1 - \frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{\lambda_2}{\lambda_1 + \lambda_2}.$$

Thus,

$$X \succeq_{SP} Y \Leftrightarrow Q(X, Y) \geq \frac{1}{2} \Leftrightarrow \frac{\lambda_2}{\lambda_1 + \lambda_2} \geq \frac{1}{2} \Leftrightarrow \lambda_2 \geq \lambda_1. \qquad \blacksquare$$

Remark 3.69 *Thevalue of the probabilist ic relation Q for independent and exponen- tial ly distributed random variables was already studied in [56, Section 6.2.1]. However, insuch reference theauthors made a mistake duringthe computations and found an incorrect expression forthe probabilisticrelation.*

As with Bernoulli distributed random variables, statistical preference and sto chastic dominance are equivalent prop erties for exponential distributions. In thiscase, also first degree sto chastic dominance, and therefore the n-degree sto chastic dominance, are complete relations, and the can be reduced to the comparison of the exp ectations.

Corollary 3.70 *Let X and Y be two independent random variables with exponential distribution. Then,*

$$X \quad {}_{FSD} \; Y \qquad X \quad {}_{nSD} \; Y \text{ for any } n \geq 2 \qquad X \quad {}_{SP} \; Y \qquad E[X] \geq E[Y].$$

Next we focus on uniform distributions.

Proposition 3.71 *Let X and Y be two independent random variables with uniform distributions, $U(a, b)$ and $U(c, d)$ respectively.*

1. *If $(a, b) \quad (c, d)$ then:*
   - $Q(X, Y) = \frac{2b-c-d}{2(b-a)}$ *and*
   - $X \quad {}_{SP} \; Y$ *if and only if* $a+b \geq c+d$.
2. *If $c \leq a < d \leq b$, X is always statistically preferred to Y, and its degree of preference is $Q(X, Y) = 1 - \frac{(d-a)^2}{2(b-a)(d-c)}$.*

Proof

1. Suppose that $a \leq c < d \leq b$. Then,

$$Q(X, Y) = P(X > Y) = \int_d^b \frac{1}{b-a} dx + \int_c^d \int_c^x \frac{1}{b-a} \frac{1}{d-c} dy\, dx$$
$$= \frac{b-d}{b-a} + \int_c^d \frac{1}{b-a} \frac{x-c}{d-c} dx = \frac{b-d}{b-a} + \frac{(d-c)^2}{2(d-c)(b-a)} = \frac{2b-c-d}{2(b-a)}.$$

   Then, $X \quad {}_{SP} \; Y$ if and only if:

$$\frac{2b-c-d}{2(b-a)} \geq \frac{1}{2} \qquad b+a \geq c+d.$$

   If $c \leq a < b \leq d$, we can similarly see that

$$Q(X, Y) = \frac{b+a-2c}{2(d-c)}.$$

   Thus, $Q(X, Y) \geq \frac{1}{2}$ if and only if $a+b \geq c+d$.

2. If $c \leq a < d \leq b$, it is easy to prove that $X \quad {}_{FSD} \; Y$, and therefore $X \quad {}_{SP} \; Y$. Let us now compute the preference degree:

$$P(Y > X) = \int_a^d \int_a^y \frac{dx\, dy}{(b-a)(d-c)} = \int_a^d \frac{y-a}{(b-a)(d-c)} dy = \frac{(d-a)^2}{2(b-a)(d-c)}.$$

   Then, $Q(X, Y) = 1 - Q(Y, X) = 1 - P(Y > X) = 1 - \frac{(d-a)^2}{2(b-a)(d-c)}$. ∎

Remark 3.72 *The value of the probabilistic relation Q for the uniform distribution was already studied in [56]. However, the authors only focused on uniform distribution with a fixed amplitude of the support, and the only parameter was the starting point of the support. This is a particular case included in the last result, and in that case, as we have seen, the random variable with the greatest minimum of the support stochastically dominates the other one, and consequently it is also statistically preferred.*

For uniform distributions, first degree stochastic dominance and statistical preference are not equivalent in general. In fact, first degree stochastic dominance does not hold when the first case of the proof of the previous proposition holds. Nevertheless, we can establish the following:

Corollary 3.73 *Let X and Y be two independent random variables with uniform distribution. It holds that:*

$$X \succeq_{FSD} Y \Leftrightarrow X \succeq_{SP} Y \Leftrightarrow E[X] \geq E[Y].$$

We next focus on the family of Pareto distribution.

Proposition 3.74 *Let X and Y be two independent random variables with Pareto distributions, $X \equiv P_a(\lambda_1)$ and $Y \equiv P_a(\lambda_2)$, respectively. Then:*

- *$Q(X, Y) = \frac{\lambda_2}{\lambda_1+\lambda_2}$ and*
- *X is statistically preferred to Y if and only if $\lambda_2 \geq \lambda_1$.*

Proof First of all, let us determine the expression of Q:

$$\begin{aligned} Q(X, Y) = P(X > Y) &= \int_1^{\infty} \int_1^{x} \lambda_1 x^{-\lambda_1-1} \lambda_2 y^{-\lambda_2-1} dy\, dx \\ &= \int_1^{\infty} \lambda_1 x^{-\lambda_1-1} \left(1 - x^{-\lambda_2}\right) dx = 1 - \frac{\lambda_1}{\lambda_1+\lambda_2}. \end{aligned}$$

Then,

$$X \succeq_{SP} Y \Leftrightarrow 1 - \frac{\lambda_1}{\lambda_1+\lambda_2} \geq \frac{1}{2} \Leftrightarrow \lambda_2 \geq \lambda_1. \qquad \blacksquare$$

As for exponential and Bernoulli distributions, the equivalence between first degree stochastic dominance and statistical preference holds for Pareto distributions. In fact, when the expectation of the random variables exists, first degree stochastic dominance is equivalent to the comparison of the expectations. Hence, it is a complete relation, and then n-th degree stochastic dominance is also complete and equivalent to first degree stochastic dominance.

Corollary 3.75 *Let X and Y be two independent random variables with Pareto distributions. Then:*

$$X \quad {}_{FSD}\ Y \qquad X \quad {}_{nSD}\ Y \text{ for any } n \geq 2 \qquad X \quad {}_{SP}\ Y.$$

Furthermore, if the parameter of X and Y are greater than 1, their expectation exists, and in that case:

$$X \quad {}_{FSD}\ Y \qquad X \quad {}_{nSD}\ Y \text{ for any } n \geq 2 \qquad X \quad {}_{SP}\ Y \qquad E[X] \geq E[Y].$$

Concerning the beta distribution, we recall that its density function is given by

$$f(x) = \begin{cases} \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} x^{p-1}(1-x)^{q-1} & \text{if } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.17}$$

However, the results we investigate in this section fix the value of one of the parameters to 1. We start by fixing $q = 1$. We obtain the following:

Proposition 3.76 *Let X and Y be two independent random variables with beta distributions, $X \equiv \beta(p_1, 1)$ and $Y \equiv \beta(p_2, 1)$, respectively. Then:*

- *$Q(X, Y) = \frac{p_1}{p_1 + p_2}$ and*
- *$X \quad {}_{SP}\ Y$ if and only if $p_1 \geq p_2$.*

Proof We first compute the expression of the relation Q.

$$Q(X, Y) = P(X > Y) = \int_0^1 \int_0^x p_1 x^{p_1 - 1} p_2 y^{p_2 - 1} dy\, dx = \int_0^1 p_1 x^{p_1 - 1} x^{p_2} dx = \frac{p_1}{p_1 + p_2}.$$

Then it holds that

$$X \quad {}_{SP}\ Y \qquad \frac{p_1}{p_1 + p_2} \geq \frac{1}{2} \qquad p_1 \geq p_2. \qquad \blacksquare$$

Taking into account that the expectation of a beta distribution with parameter $q = 1$ is $\frac{p}{p+1}$, the equivalence between statistical preference and the comparison of the expectations is clear. Furthermore, take into account that the cumulative distribution function associated with a beta distribution with parameter $q = 1$ is given by:

$$F(x) = \begin{cases} 0 & \text{if } x \leq 0. \\ x^p & \text{if } 0 < x < 1. \\ 1 & \text{if } x \geq 1. \end{cases}$$

Then, it is clear that sto chastic dominance between two variables of this typ e can be reduced to verifying which of the parameters p is greater. Finally, itis easytocheck that this is equivalent to take the variable with greater exp ectation. Thus, inthis case sto chastic dominance, statistical preference and the comparison of exp ectations are also equivalent.

Corollary 3.77 *Let X and Y be two independent randomvariables with beta distributions with second paramet er equal to 1. Then,*

$$X \quad {}_{FSD}\ Y \quad X \quad {}_{nSD}\ Y \text{ for any } n \geq 2 \quad X \quad {}_{SP}\ Y \quad E[X] \geq E[Y].$$

Finally, we consider b eta distributions with $p=1$.

Prop osition 3.78 *Let X and Y betwoindependentrandomvariableswith distributions $X \equiv \beta(1, q_1)$ and $Y \equiv \beta(1, q_2)$, respectively. Then:*

- $Q(X, Y) = \frac{q_2}{q_1+q_2}$ *and*
- $X \quad {}_{SP}\ Y$ *if andonly if* $q_2 \geq q_1$.

Pro of In order to prove the result, note that $X \equiv \beta(1, q) \quad 1 - X \equiv \beta(q, 1)$

$$F_{1-X}(t) = P(1 - X \leq t) = 1 - F_X(1-t) = 1 - [1 - (1 - (1-t))^q] = t^q.$$

Then, taking into account Prop osition 3.3, $X \quad {}_{SP}\ Y \quad 1 - Y \quad {}_{SP}\ 1 - X$ and $Q(X, Y) = Q(1 - Y, 1 - X) = \frac{q_2}{q_1+q_2}$, and using Prop osition 3.76, statistical preference is equivalent to $q_2 \geq q_1$. ■

As in the previous case, since the exp ectation of a b eta distribution with parameter $p=1$ is $\frac{1}{1+q}$, the equivalence b etween sto chastic dominance and statistical preference also holds for beta distributions.

Corollary 3.79 *Let X and Y be two independent randomvariables with beta distributions withfirst parameter equal to 1. Then,*

$$X \quad {}_{FSD}\ Y \quad X \quad {}_{nSD}\ Y \text{ for any } n \geq 2 \quad X \quad {}_{SP}\ Y \quad E[X] \geq E[Y].$$

The normal distribution

We now study normally distributed random variables. Inthiscasewe will not only consider indep endent variables. Thus, we b egin with the comparison of one-dime nsional distributions and then we shall consider the case of the comparison of the comp onents of a bidime nsional random vector normally distributed.

Proposition 3.80 *Let X and Y be two independent and normally distributed random variables, $N(\mu_1,\sigma_1)$ and $N(\mu_2,\sigma_2)$, respectively. Then, X will be statistically preferred to Y if and only if $\mu_1 \geq \mu_2$.*

Proof The relation Q takes the value (see [56, Section 7]):

$$Q(X,Y) = F_{N(0,1)} \frac{\mu_1 - \mu_2}{\sigma_1^2 + \sigma_2^2} .$$

Then:

$$X \; {}_{SP} \; Y \quad Q(X,Y) \geq \frac{1}{2} \quad F_{N(0,1)} \frac{\mu_1 - \mu_2}{\sigma_1^2 + \sigma_2^2} \geq \frac{1}{2} \quad \frac{\mu_1 - \mu_2}{\sigma_1^2 + \sigma_2^2} \geq 0 \quad \mu_1 \geq \mu_2 .$$

∎

Given two normally distributed random variables X $N(\mu_1,\sigma_1)$ and Y $N(\mu_2,\sigma_2)$, it holds that $X \; {}_{FSD} \; Y$ if and only if they are identically distributed, $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$, (see [139]). Then, statistical preference is not equivalent to first degree stochastic dominance for normal random variables.

For independent normal distributions, the variance of the variables are not important when studying statistical preference. For this reason, statistical preference is equivalent to the criterium of maximum mean in the comparison of normal random variables:

Corollary 3.81 *Consider two independent random variables X and Y normally distributed. It holds that:*

$$X \; {}_{FSD} \; Y \quad X \; {}_{SP} \; Y \quad E[X] \geq E[Y].$$

Let us now consider a bidimensional random vector with normal distribution:

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \equiv N \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix} \right) . \tag{3.18}$$

Now, our aim is to compare the components X_1 and X_2 of this random vector. We obtain the following result:

Theorem 3.82 *Consider the random vector $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ normally distributed as in Equation (3.18). Then, it holds that:*

- $Q(X_1, X_2) = F_{N(0,1)} \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}}$.

- $X_1 \quad {}_{SP} X_2 \quad \mu_1 \geq \mu_2$.

Pro of Applyingthe usual prop erties of the normal distributions, thedistribution of $X_1 - X_2$ is:

$$X_1 - X_2 = (1 \quad -1) \begin{matrix} X_1 \\ X_2 \end{matrix} \equiv N\left((1 \quad -1) \begin{matrix} \mu_1 \\ \mu_2 \end{matrix}, (1 \quad -1) \begin{matrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{matrix} \begin{matrix} 1 \\ -1 \end{matrix} \right)$$
$$= N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2),$$

where the second parameter is consider to be the variance instead of the standard deviation. Then:

$$P(X_1 > X_2) = P(X_1 - X_2 > 0) = P\left(N(0,1) > \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}} \right)$$
$$= P\left(N(0,1) < \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}} \right) = F_{N(0,1)}\left(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}} \right).$$

Thus, $X_1 \quad {}_{SP} X_2$ if and on ly if $F_{N(0,1)}\left(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}} \right) \geq \frac{1}{2}$. ∎

This result is more general than Prop osition 3.80, which corresp onds to the case $\rho = 0$. Moreover, in that case statistic al preference is also equivalent to the comparison of the exp ectations. However, theadvantageofobtaining a degreeof preference isobvious. In fact, we haveto recall the influence of the correlation co efficient ρ in the value of the preference degree: although the preference between X_1 and X_2 is only basedon the comparison of the exp ectations($X_1 \quad {}_{SP} X_2 \quad \mu_1 \geq \mu_2$), the valueof ρ plays an imp ortant role for the preference degree. For instance, the greater the correlati on co efficient,the greater thepreference degree $Q(X, Y)$. For thisreason, thegreater the correlation co efficient, the stronger the preference of X over Y.

In Table 3.4 we have summarised the res ults that we have obtained in this subsection.

As a summary, we have seen that for the some of usual distributions in indep endent random variables, statistical preference is equivalent to the comparison of its exp ectations, andin several cases,sto chastic dominance and statistical preference are also equivalent. Let usrecall that, in particular, for thedistributionswehavestudiedthat b elongs to the exponential family of distribu tions, sto chastic dominance and statistical preference are equivalent. We can conjecture that for indep endent random variables whose distribution b elong to the exp onential family of distributions, statistical preference and sto chastic dominance are equivalent, and are also equivalent to the comparison of the exp ectations.

Nevertheless, at this p oint, this is just a c on jecture b ecause it has not b een proved yet.

Distributions	$Q(X_1, X_2)$	Condition
$X_i \equiv B(p_i)$, $i = 1,2$	$\frac{1}{2}(p_1 - p_2 + 1)$	$p_1 \geq p_2$
$X_i \equiv Exp(\lambda_i)$, $i = 1,2$	$\frac{\lambda_2}{\lambda_1+\lambda_2}$	$\lambda_2 \geq \lambda_1$
$X_1 \equiv U(a,b)$, $X_2 \equiv U(c,d)$		
$a \leq c \leq d < b$	$\frac{2b-c-d}{2(b-a)}$	$a+b \geq c+d$
$c < a < b \leq d$	$\frac{a+b-2c}{2(d-c)}$	$a+b \geq c+d$
$c \leq a < d \leq b$	$1 - \frac{(d-a)^2}{2(d-c)(b-a)}$	Always
$Pa(\lambda_i)$, $i = 1,2$	$\frac{\lambda_2}{\lambda_1+\lambda_2}$	$\lambda_2 \geq \lambda_1$
$\beta(p_i, 1)$, $i = 1,2$	$\frac{p_1}{p_1+p_2}$	$p_1 \geq p_2$
$\beta(1, q_i)$, $i = 1,2$	$\frac{q_2}{q_1+q_2}$	$q_2 \geq q_1$
$N(\mu_i, \sigma_i)$, $i = 1,2$	$F_{N(0,1)}\left(\frac{\mu_1-\mu_2}{\sqrt{\sigma_1^2+\sigma_2^2}}\right)$	$\mu_1 \geq \mu_2$

Table 3.4: Characterizations of statistical preference between independent random variables included in the same family of distributions.

Although during this paragraph we have assumed independence for non-normally distributed variables, there are other cases of interest. For instance, in [32] the case of comonotonic and countermonotonic random variables are studied. In particular, Proposition 3.65, that assures that

$$X \quad _{SP} Y \qquad X \quad _{nSD} Y \qquad E[X] \geq E[Y] \text{ for any } n \geq 1$$

for independent random variables with Bernoulli distribution, could be easily extended to Bernoulli distributed random variables, taking into account the possible dependence relationship between them.

3.3 Comparison of variables by means of the statistical preference

So far, we have investigated several properties of stochastic dominance and statistical preference as pairwise comparison methods. However, a natural question arises: can we employ those methods for the comparison of more than two variables? On the one hand, stochastic dominance was defined as a pairwise comparison method, based on the direct comparison of the cumulative distribution functions, or their iterative integrals. As we already mentioned, stochastic dominance allows for incomparability. Thus, if incomparability can happen when comparing two distribution functions, it should be more frequent when comparing more than two. Then, stochastic dominance does not seem to be a good alternative for the comparison of more than two variables.

On the other hand, statistical preference has an important drawback: its lack of transitivity. The idea of statistical preference is to consider X preferred to Y when it provides greater utility the majority of times. As such, it is close to the rule of majority in voting systems; taking into account Condorcet's paradox (see [40]) it is not difficult to see that statistical preference is not transitive. When De Schuymer et al. ([55, 57]) introduced this notion, they provided an example to illustrate this fact; another one can be found in [67, Example 3].

Example 3.83 ([57, Section 1]) *As in Example 3.10, consider the following dice:*

$$A = \{1, 3, 4, 15, 16, 17\}$$
$$B = \{2, 10, 11, 12, 13, 14\} \tag{3.19}$$

and also the dice

$$C = \{5, 6, 7, 8, 9, 18\}$$

where by dice we mean a discrete and uniformly distributed random variable. We consider the game consisting on rolling the three dice simultaneously, so that the dice whose number is greater wins the game. Thus, A, B and C can be seen as independent random variables.

If we compute the probabilistic relation Q for these dices we obtain the following results:

$$Q(A, B) = \frac{5}{9} \quad A \succ_{SP} B.$$
$$Q(B, C) = \frac{25}{36} \quad B \succ_{SP} C.$$
$$Q(C, A) = \frac{7}{12} \quad C \succ_{SP} A.$$

Hence, dice A is statistically preferred to dice B, dice B is statistically preferred to dice C but dice C is statistically preferred to dice A, that is, there is a cycle, as we can see in Figure 3.4.

This fact is known as the cycle-transitivity problem, and it has already been studied by some authors, like De Shuymer et al. ([14, 15, 16, 49, 54, 56, 57, 58]) and Martinetti et al. ([122]).

This shows that statistical preference could not be adequate when we want to compare more than two random variables, precisely because it is based on pairwise comparisons.

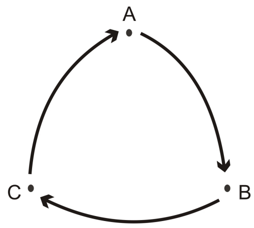


Figure 3.4: Probabilisticrelationforthe threedices.

Since b oth sto chastic dominance and statistical preference do not seem to b e adequate metho ds for the comparison of more than two variables, our aim in this section is to provide ageneralisation of the statistical preference for the comparison of n random variables, based ona extension of the probabili stic relation defined in Equation (2.7). After intro ducing the main definition, we shall investigate its prop erties, its possible characterizationsanditsconnectionwiththe"usual" statisticalpreference, aswellasits possible relationships with sto chastic dominance.

3.3.1 generalisation of the statistical preference

First of all we are goingto analyze the case of three random variables, as in the dice example, and later weshall generalise our definition to the case of n random variables.

Let us consider three random variables denoted by X, Y and Z defined onthe probability sp ace$(\Omega, \mathcal{A}, P)$. We can decomp ose Ω in the following way:

$$\begin{aligned}\Omega = & \{X > \max(Y, Z)\} \ \{Y > \max(X, Z)\} \ \{Z > \max(X, Y)\} \\ & \{X = Y > Z\} \ \{X = Z > Y\} \ \{Y = Z > X\} \ \{X = Y = Z\}\end{aligned} \tag{3.20}$$

Obviously, $\{X > \max(Y, Z)\}$ denotes the subset of Ω formed by the elements ω Ω satisfying $X(\omega) > \max(Y(\omega), Z(\omega))$, andsimilarlyfor theothers. In what remains we will use the short wayin order to simplify the notation.

This is a decomp osition of Ω into pairwise disjointsubsets, i.e., a partitionof Ω. As aconsequence,

$$\begin{aligned}1 = & P(X > \max(Y, Z)) + P(Y > \max(X, Z)) + P(Z > \max(X, Y)) + P(X = Y > Z) \\ & + P(X = Z > Y) + P(Y = Z > X) + P(X = Y = Z).\end{aligned} \tag{3.21}$$

Since our goal is to define the degree in which X is preferred to Y and Z, we can define $Q_2(X, [Y, Z])$ by the following equation:

$$Q_2(X, [Y, Z]) = P(X > \max(Y, Z)) + \frac{1}{2}\left(P(X = Y > Z) + P(X = Z > Y)\right) + \frac{1}{3}P(X = Y = Z).$$

This generalises Equation (2.7). Furthermore, if we consider $Q_2(Y, [X, Z])$ and $Q_2(Z, [X, Y])$ given by:

$$Q_2(Y, [X, Z]) = P(Y > \max(X, Z)) + \frac{1}{2}\left(P(X = Y > Z) + P(Y = Z > X)\right) + \frac{1}{3}P(X = Y = Z);$$

$$Q_2(Z, [X, Y]) = P(Z > \max(X, Y)) + \frac{1}{2}\left(P(X = Z > Y) + P(Y = Z > X)\right) + \frac{1}{3}P(X = Y = Z);$$

using the partition of Ω showed in Equation (3.20) and Equation (3.21), it can be shown that:

$$Q_2(X, [Y, Z]) + Q_2(Y, [X, Z]) + Q_2(Z, [X, Y]) = 1.$$

In this sense, following the idea of De Schuymer et al. ([55, 57]), X can be considered preferred to Y and Z if

$$Q_2(X, [Y, Z]) \geq \max\{Q_2(Y, [X, Z]), Q_2(Z, [X, Y])\}.$$

Moreover, X is preferred to Y and Z with degree $Q_2(X, [Y, Z])$.

More generally, we can consider a set of alternatives D formed by some random variables defined on the same probability space. Then, we can consider the map:

$$Q_n : D \times D^n \to [0, 1],$$

defined by:

$$\begin{aligned} Q_n(X, [X_1, \ldots, X_n]) = {} & \operatorname{Prob}\{X > \max(X_1, \ldots, X_n)\} \\ & + \frac{1}{2}\sum_{i=1}^{n} \operatorname{Prob}\{X = X_i > \max(X_j : j \neq i)\} \\ & + \frac{1}{3}\sum_{1 \leq i < j \leq n} \operatorname{Prob}\{X = X_i = X_j > \max(X_k : k \neq i, j)\} \\ & + \ldots + \frac{1}{n+1}\operatorname{Prob}\{X = X_1 = \ldots = X_n\}. \end{aligned}$$

Equivalently, the relation Q_n can be expressed by:

$$Q_n(X, [X_1, \ldots, X_n]) = \sum_{\substack{k = 0, \ldots, n \\ 1 \leq i_1 < \ldots < i_k \leq n}} \frac{1}{k+1} P\left(X = X_{i_1} = \ldots = X_{i_k} > \max_{j \neq i_1, \ldots, i_k}(X_j)\right), \quad (3.22)$$

where $\{i_1, \ldots, i_k\}$ denotes any ordered subset of k-elements of $\{1, \ldots, n\}$. Note that this formula is the generalisation of the probabilistic relation defined on Equation (2.7), since for $n=1$ we obtain the expression of such probabilistic relation. We can interpret the value of $Q_n(X, [X_1, \ldots, X_n])$ as the degree in which X is preferred to $X_1, \ldots, X_n$. Consequently, the greater the value of $Q_n(X, [X_1, \ldots, X_n])$ the stronger the preference of X over $X_1, \ldots, X_n$. The relation Q_n allows to define the concept of general statistical preference.

Definition 3.84 *Let $X, X_1, \ldots, X_n$ be $n+1$ random variables. X is statistically preferred to $X_1, \ldots, X_n$, and it is denoted by $X \succeq_{SP} [X_1, \ldots, X_n]$, if*

$$Q_n(X, [X_1, \ldots, X_n]) \geq \max_{i=1,\ldots,n} Q_n(X_i, [X, \{X_j : j \neq i\}]). \tag{3.23}$$

As it was the case for statistical preference, this generalisation uses the joint distribution of the variables, and thus takes into account the stochastic dependencies between them. Moreover, the relation Q_n provides a degree of preference of a random variable with respect to the others, and through this we can establish which is the preferred random variable, the second preferred random variable, etc. For instance, if $Q_n(X_i, [X, \{X_j : j \neq i\}]) \geq Q_n(X_j, [X, \{X_j : j \neq i\}])$ for every $i > j$ and Equation (3.23) holds, then X is the preferred random variable, X_1 is the second preferred random variable and, in general, X_i is the $i+1$ preferred random variable, with their respective degrees of preference.

Example 3.85 *If we consider the dices defined on Equation (3.19) and apply the general statistical preference to find the preferred dice, we obtain the following preference degrees: $Q_2(X, [Y, Z]) = 0.4167$, $Q_2(Y, [X, Z]) = 0.3472$ and $Q_2(Z, [X, Y]) = 0.2361$. Thus, X is the preferred dice with degree 0.4167, Y is the second preferred dice with degree 0.3472; and Z is the less preferred dice with degree 0.2361.*

3.3.2 Basic properties

In this subsection we investigate some basic properties of the general statistical preference. The first part is devoted to the study of the relationships between pairwise statistical preference and general preference. Similarly, we also establish a connection between $Q(\cdot, \cdot)$ and $Q_n(\cdot, [\cdot])$. Finally, we generalise Proposition 3.39 and Theorem 3.40, where we showed the connection between statistical preference and the median, for the general statistical preference and establish a characterization of this notion.

Consider random variables $X, X_1, \ldots, X_n$. In our first result we prove that general statistical preference sometimes offers a different preferred random variable than pairwise statistical preference. This is because general statistical preference uses the joint distribution of all the variables, while pairwise statistical preference only takes into account their bivariate distributions, and consequently it does not use all the available information.

Proposition 3.86 *Let $X, X_1, \ldots, X_n$ be $n+1$ random variables. It holds that:*

- *There are $X, X_1, \ldots, X_n$ random variables such that $X \ {}_{SP}\ X_i$ for every $i = 1, \ldots, n$ and $X_j \ {}_{SP}\ [X, X_i : i = j]$ for some $j \ \{1, \ldots, n\}$.*
- *There are $X, X_1, \ldots, X_n$ random variables such that $X_i \ {}_{SP}\ X$ for every $i = 1, \ldots, n$ and $X \ {}_{SP}\ [X_1, \ldots, X_n]$.*

Proof Let us consider the first statement. To see that the implication does not hold in general, consider $n=2$ and the independent random variables X, X_1 and X_2 defined by:

X	3	5
P_X	0.5	0.5

X_1	0	5
P_{X_1}	0.5	0.5

X_2	2	6
P_{X_2}	0.51	0.49

For these variables it holds that $Q(X, X_1) = 0.625$ and $Q(X, X_2) = 0.51$, and consequently $X \ {}_{SP}\ X_1$ and $X \ {}_{SP}\ X_2$. However,

$$
\begin{aligned}
Q_2(X, [X_1, X_2]) &= 0.31875. \\
Q_2(X_1, [X, X_2]) &= 0.19125. \\
Q_2(X_2, [X, X_1]) &= 0.49.
\end{aligned}
$$

Thus, $X_2 \ {}_{SP}\ [X, X_1]$.

Consider now the second statement. Consider $n=2$ and the independent dices X, X_1 and X_2 defined by:

$$
\begin{aligned}
X &= \{1, 2, 4, 6, 17, 18\} \\
X_1 &= \{3, 7, 9, 12, 14, 16\} \\
X_2 &= \{5, 8, 10, 11, 13, 15\}
\end{aligned}
$$

It holds that $X_1 \ {}_{SP}\ X$ and $X_2 \ {}_{SP}\ X$, since $Q(X, X_1) = \frac{7}{18}$ and $Q(X, X_2) = \frac{13}{36}$. However, if we compute the relation $Q_2(\ ,[\ ])$ we obtain the following:

$$
\begin{aligned}
Q_2(X, [X_1, X_2]) &= \tfrac{73}{216}. \\
Q_2(X_1, [X, X_2]) &= \tfrac{72}{216}. \\
Q_2(X_2, [X, X_1]) &= \tfrac{71}{216}.
\end{aligned}
$$

Consequently, $X \ {}_{SP}\ [X_1, X_2]$. ∎

Next we prove that $Q_n(X, [X_1, \ldots, X_n])$ is always lower than or equal to $Q(X, X_i)$.

Proposition 3.87 *Let us consider the random variables $X, X_1, \ldots, X_n$. It holds that:*

$$Q_n(X, [X_1, \ldots, X_n]) \leq Q(X, X_i) \text{ for every } i = 1, \ldots, n.$$

Consequently, if $Q_n(X, [X_1, \ldots, X_n]) \geq \frac{1}{2}$, then $X \ {}_{SP}\ [X_1, \ldots, X_n]$ and $X \ {}_{SP}\ X_i$ for every $i = 1, \ldots, n$.

Proof Recall that $Q(X, X_i) = P(X > X_i) + \frac{1}{2} P(X = X_i)$. It holds that:

$$\{X > X_i\} \supseteq \bigcup_{\substack{k=0,\dots,n-1 \\ i_1,\dots,i_k \neq i}} \left\{ X = X_{i_1} = \dots = X_{i_k} > \max_{j \neq i, i_1,\dots,i_k} (X_i, X_j) \right\}.$$

Moreover, the previous sets are pairwise disjoint, and consequently:

$$P(X > X_i) \geq \sum_{\substack{k=0,\dots,n-1 \\ i_1,\dots,i_k \neq i}} P\left(X = X_{i_1} = \dots = X_{i_k} > \max_{j \neq i, i_1,\dots,i_k} (X_i, X_j) \right).$$

Similarly:

$$\{X = X_i\} \supseteq \bigcup_{\substack{k=0,\dots,n-1 \\ i_1,\dots,i_k \neq i}} \left\{ X = X_i = X_{i_1} = \dots = X_{i_k} > \max_{j \neq i, i_1,\dots,i_k} (X_j) \right\}.$$

Since these sets are pairwise disjoint,

$$P(X = X_i) \geq \sum_{\substack{k=0,\dots,n-1 \\ i_1,\dots,i_k \neq i}} P\left(X = X_i = X_{i_1} = \dots = X_{i_k} > \max_{j \neq i_1,\dots,i_k} (X_j) \right).$$

Consequently, we obtain that:

$$\begin{aligned}
Q(X, X_i) = P(X > X_i) + \tfrac{1}{2} P(X = X_i) \geq \\
\sum_{\substack{k=0,\dots,n-1 \\ i_1,\dots,i_k \neq i}} P\left(X = X_{i_1} = \dots = X_{i_k} > \max_{j \neq i, i_1,\dots,i_k} (X_i, X_j)\right) + \\
\tfrac{1}{2} \sum_{\substack{k=0,\dots,n-1 \\ i_1,\dots,i_k \neq i}} P\left(X = X_i = X_{i_1} = \dots = X_{i_k} > \max_{j \neq i_1,\dots,i_k} (X_j)\right) \geq \\
\sum_{\substack{k=0,\dots,n \\ i_1,\dots,i_k \in \{1,\dots,n\}}} \frac{1}{k+1} P\left(X = X_{i_1} = \dots = X_{i_k} > \max_{j \neq i_1,\dots,i_k} (X_j)\right) = \\
Q_n(X, [X_1, \dots, X_n]).
\end{aligned}$$

We conclude that $Q(X, X_i) \geq Q_n(X, [X_1, \dots, X_n])$. Consequently, if

$$Q_n(X, [X_1, \dots, X_n]) \geq \frac{1}{2}$$

then $X \succeq_{SP} [X_1, \dots, X_n]$ and $X \succeq_{SP} X_i$ for every $i = 1, \dots, n$. ∎

Next we establish the connection between the probabilistic relation $Q(\cdot, \cdot)$ and $Q_n(\cdot, [\cdot])$.

Proposition 3.88 *Let $X, X_1, \ldots, X_n$ be $n+1$ random variables defined on the same probability space. It holds that*

$$Q_n(X, [X_1, \ldots, X_n]) - Q(X, \max(X_1, \ldots, X_n)) = \sum_{k=2}^{n} \left(\frac{1}{k+1} - \frac{1}{2}\right) \sum_{\substack{1 \leq i_1 < \ldots < i_k \leq n \\ i_j \neq i_l \, j \neq l}} P\left(X = X_{i_1} = \ldots = X_{i_k} > \max_{l \neq i_1, \ldots, i_k} (X_l)\right).$$

Proof Consider the expression of $Q(X, \max(X_1, \ldots, X_n))$:

$$Q(X, \max(X_1, \ldots, X_n)) = P(X > \max(X_1, \ldots, X_n)) + \frac{1}{2} \sum_{k=1}^{n} \sum_{\substack{1 \leq i_1 < \ldots < i_k \leq n \\ i_j \neq i_l \, j \neq l}} P\left(X = X_{i_1} = \ldots = X_{i_k} > \max_{l \neq i_1, \ldots, i_k} (X_l)\right).$$

Using Equation (3.22), $Q_n(X, [X_1, \ldots, X_n])$ can be expressed by:

$$Q_n(X, [X_1, \ldots, X_n]) = P(X > \max(X_1, \ldots, X_n)) + \sum_{k=1}^{n} \frac{1}{k+1} \sum_{\substack{1 \leq i_1 < \ldots < i_k \leq n \\ i_j \neq i_l \, j \neq l}} P\left(X = X_{i_1} = \ldots = X_{i_k} > \max_{l \neq i_1, \ldots, i_k} (X_l)\right).$$

The result follows simply by making the difference between both expressions. ∎

From this result we deduce that

$$Q_n(X, [X_1, \ldots, X_n]) \leq Q(X, \max(X_1, \ldots, X_n)). \tag{3.24}$$

Then, if $X \succeq_{SP} [X_1, \ldots, X_n]$ holds with degree $Q_n(X, [X_1, \ldots, X_n]) \geq \frac{1}{2}$, we obtain $X \succeq_{SP} \max(X_1, \ldots, X_n)$.

Moreover, there are situations where the inequality of Equation (3.24) becomes an equality. To see this, let us introduce the following notation:

$$X_{-i} = \{X_j : j \neq i\}.$$

Corollary 3.89 *Under the conditions of the previous proposition, if for every $k \in \{1, \ldots, n\}$ and for every $1 \leq i_1 < \ldots < i_k$ it holds that*

$$P(X = X_{i_1} = \ldots = X_{i_k} > \max(X_j : j \neq i_1, \ldots, i_k)) = 0, \tag{3.25}$$

then

$$Q_n(X, [X_1, \ldots, X_n]) = Q(X, \max(X_1, \ldots, X_n)).$$

Furthermore, if for every k $\{1, \dots, n\}$ *and for every* $1 \leq i_1 < \dots < i_k \leq n$ *it holds that*

$$P(X_{i_1} = \dots = X_{i_k} > \max(X, X_j : j = i_1, \dots, i_k)) = 0, \qquad (3.26)$$

then

$$Q_n(X_i, [X, X_{-i}]) = Q(X_i, \max(X, X_{-i})),$$

for every $i = 1, \dots, n$.

In particular, the previous result holds when the random variables satisfy, $P(X = X_i) = P(X = X_j) = 0$ for every $i = j$, as is for instance the case with discrete random variables with pairwise disjoint supports.

Finally, let us generalise Theorem 3.40 and to provide a characterization of general statistical preference. For this aim we consider random variables $X, X_1, \dots, X_n$ satisfying Equations (3.25) and (3.26) for every k $\{0, \dots, n\}$ and every $1 \leq 1_i < \dots < i_k \leq n$. Although this restriction will be imposed also in Theorems 3.91, 3.95 and Lemma 3.94, it is not too restrictive. In fact, it is satisfied by discrete random variables with pairwise disjoint supports or absolutely continuous random vectors $(X, X_1, \dots, X_n)$. Consequently, we can understand it as a technical condition.

Theorem 3.90 *Let* $X, X_1, \dots, X_n$ *be* $n+1$ *real-valued random variables defined on the same probability satisfying Equations* (3.25) *and* (3.26). *Then,* $X \ {}_{SP} [X_1, \dots, X_n]$ *holds if and only if*

$$F_{X - \max(X_1, \dots, X_n)}(0) \leq F_{X_i - \max(X, X_{-i})}(0) \text{ for every } i = 1, \dots, n.$$

Proof The probabilistic relation $Q(X, Y)$ can be expressed by:

$$Q(X, Y) = 1 - F_{X-Y}(0) + \frac{1}{2} P(X = Y).$$

Thus, using this expression and applying Corollary 3.89 it holds that:

$$\begin{aligned} Q_n(X, [X_1, \dots, X_n]) = Q(X, \max(X_1, \dots, X_n)) &= 1 - F_{X - \max(X_1, \dots, X_n)}(0) \\ + \frac{1}{2} P(X = \max(X_1, \dots, X_n)) &= 1 - F_{X - \max(X_1, \dots, X_n)}(0). \end{aligned}$$

Similarly, we can compute the value of $Q_n(X_i, [X, X_{-i}])$:

$$Q_n(X_i, [X, X_{-i}]) = 1 - F_{X_i - \max(X, X_{-i})}(0).$$

Therefore, $X \ {}_{SP} [X_1, \dots, X_n]$ if and only if:

$$1 - F_{X - \max(X_1, \dots, X_n)}(0) \geq 1 - F_{X_i - \max(X, X_{-i})}(0),$$

or equivalently,

$$F_{X-\max(X_1,\ldots,X_n)}(0) \leq F_{X_i-\max(X,X_{-i})}(0)$$

for every $i = 1,\ldots,n$. ∎

Thus, given random variables $X, X_1,\ldots,X_n$ in the conditions of the previous result, to find the preferred one by computing the values of $Q_n(\ ,[\ ])$ is equivalent to comparing the values of $F_{X-\max(X_1,\ldots,X_n)}(0)$ and $F_{X_i-\max(X,X_{-i})}(0)$ for $i = 1,\ldots,n$.

3.3.3 Stochastic dominance Vs general statistical preference

In Section 3.2 we saw that in a number of cases first degree stochastic dominance implies statistical preference for real-valued random variables. Now we investigate the connection between stochastic dominance and general statistical preference. Again, we shall consider different cases: on the one hand, independent and comonotonic random variables, for which we shall obtain an equivalent expression for $Q_n(\ ,[\ ])$. On the other hand, we shall consider random variables coupled by Archimedean copulas. Recall that we omit countermonotonic random variables since, as we already said, the Łukasiewicz operator is not a copula for $n \geq 2$. Finally, we also investigate the relationships between the n^{th} degree stochastic dominance and general statistical preference.

Independent and comonotonic random variables

Let us begin our study with the case of independent real-valued random variables. In this case, by generalizing Theorem 3.44, we deduce that first degree stochastic dominance implies general statistical preference.

Theorem 3.91 *Let us consider $X, X_1,\ldots,X_n$ independent real-valued random variables satisfying Equations* (3.25) *and* (3.26). *Then, if $X \;_{FSD}\; X_i$ for $i = 1,\ldots,n$, implies $X \;_{SP}\; [X_1,\ldots,X_n]$.*

Proof Since we are under the hypotheses of Corollary 3.89, we deduce that:

$$Q_n(X,[X_1,\ldots,X_n]) = Q(X,\max(X_1,\ldots,X_n)) \text{ and}$$
$$Q_n(X_i,[X,X_{-i}]) = Q(X_i,\max(X,X_{-i})),$$

for every $i = 1,\ldots,n$. Therefore, $X \;_{SP}\; [X_1,\ldots,X_n]$ if and only if:

$$P(X \geq \max(X_1,\ldots,X_n)) \geq P(X_i \geq \max(X,X_{-i})), \quad i = 1,\ldots,n.$$

Note that, since $X, X_1,\ldots,X_n$ are independent, we also have that:

- X and $\max(X_1,\ldots,X_n)$ are independent.

- X_i and $\max(X, X_{-i})$ are independent.

Now, we have to remark that, if U_1 and U_2 are two independent random variables with respective cumulative distribution functions F_{U_1} and F_{U_2}, Lemma 3.11 assures that $P\{U_1 \geq U_2\} = E[F_{U_2}(U_1)]$.

Applying this result, we deduce that:

$$P(X \geq \max(X_1, \ldots, X_n)) = E(F_{\max(X_1,\ldots,X_n)}(X)) = E(F_{X_1}(X) \ldots F_{X_n}(X)).$$

Similarly,

$$\begin{aligned} P(X_i \geq \max(X, X_{-i})) &= E(F_{\max(X,X_{-i})}(X_i)) \\ &= E[F_X(X_i) \prod_{j \neq i} F_{X_j}(X_i)] \leq E[\prod_{j=1}^{n} F_{X_j}(X_i)], \end{aligned}$$

where last inequality holds since $F_X \leq F_{X_i}$. Finally, since $X \succeq_{FSD} X_i$, Equation (2.6) assures that $E[h(X)] \geq E[h(X_i)]$ for any increasing function h. In particular, we may consider the increasing function

$$h(t) = \prod_{j=1}^{n} F_{X_j}(t).$$

Therefore,

$$\begin{aligned} P(X \geq \max(X_1, \ldots, X_n)) &= E(F_{X_1}(X) \ldots F_{X_n}(X)) \\ &\geq E[\prod_{j=1}^{n} F_{X_j}(X_i)] \geq P(X_i \geq \max(X, X_{-i})), \end{aligned}$$

or equivalently,

$$Q(X, \max(X_1, \ldots, X_n)) \geq Q(X_i, \max(X, X_{-i})).$$

We conclude that $X \succeq_{SP} [X_1, \ldots, X_n]$. ∎

Now we shall see that, as with statistical preference for independent random variables, strict first degree stochastic dominance also implies strict general statistical preference. For this aim, we need to establish the following lemma.

Lemma 3.92 *Consider $n+1$ independent real-valued random variables $X, X_1, \ldots, X_n$ satisfying Equations (3.25) and (3.26) such that $X \succ_{FSD} X_i$ for $i = 1, \ldots, n$. The following statements hold:*

1. *There is t such that $F_X(t) < F_{X_i}(t)$ and $F_{X_j}(t) > 0$ for any $j \neq i$.*
2. *If $P(X_i = t) = 0$ for any t satisfying the first point, then there exists an interval $[a, b]$ and $\varepsilon > 0$ such that:*

$$\prod_{j=1}^{n} F_{X_j}(t) - F_X(t) - \prod_{j \neq i} F_{X_j}(t) \geq \varepsilon > 0,$$

and $P(X_j \in [a, b]) > 0$.

Proof Let us prove the first statement. Ex-absurdo, assume that for any t such that $F_X(t) < F_{X_i}(t)$, there exist $j_1, \ldots, j_k$ such that $F_{X_{j_1}}(t) = F_{X_{j_k}}(t) = 0 < F_{X_j}(t)$ for any $j \neq j_1, \ldots, j_k$, and therefore $F_X(t) = 0$. Since the cumulative distribution functions are right-continuous, there is t^* such that $0 = F_X(t) < F_{X_i}(t)$ for any $t < t^*$ and $0 < F_X(t^*) \leq F_{X_j}(t^*)$ for any $j = 1, \ldots, n$. Then:

$$P(X = t^*) > 0, \quad P(X_{j_1} = t^*) > 0, \ldots, \quad P(X_{j_k} = t^*) > 0.$$

Hence:

$$P(X = X_{j_1} = \ldots = X_{j_k} > X_j : j \neq j_1, \ldots, j_k) \geq P(X = X_{j_1} = \ldots = X_{j_k} = t^* > X_j : j \neq j_1, \ldots, j_k) > 0,$$

and this contradicts Equation (3.25). We conclude that there exists at least t^* such that $F_X(t^*) < F_{X_i}(t^*)$ and $F_{X_j}(t^*) > 0$ for any $j \neq i$.

Let us now check the second statement. Let t^* be a point such that $F_X(t^*) < F_{X_i}(t^*)$ and $F_{X_j}(t^*) > 0$ for any $j \neq i$. Following the same steps as in Lemma 3.45 we can prove that the existence of an interval $[a, b]$ including t^* and $\delta > 0$ such that $F_{X_i}(t) - F_X(t) \geq \delta > 0$ for any $t \in [a, b]$ and $P(X_i \in [a, b]) > 0$. Furthermore, since by hypothesis $P(X_i = t) = 0$ for any $t \in [a, b]$ F_{X_i} should be strictly increasing in a subinterval $[a_1, b_1]$ of $[a, b]$

Now, consider a point t_0 in the interval $[a_1, b_1]$. Since all the F_{X_j}, for $j = 1, \ldots, n$, and F_X are right-continuous:

$$\lim_{\varepsilon \to 0} \prod_{j=1}^n F_{X_j}(t_0 + \varepsilon) = \prod_{j=1}^n F_{X_j}(t_0) > F_X(t_0) \prod_{j \neq i} F_{X_j}(t_0) = \lim_{\varepsilon \to 0} F_X(t_0 + \varepsilon) \prod_{j \neq i} F_{X_j}(t_0 + \varepsilon).$$

Then, there is $\varepsilon > 0$, and can we assume $\varepsilon \leq b_1 - t_0$, such that:

$$F_X(t_0 + \varepsilon) \prod_{j \neq i} F_{X_j}(t_0 + \varepsilon) \leq F_X(t_0) \prod_{j \neq i} F_{X_j}(t_0) + \frac{\prod_{j=1}^n F_{X_j}(t_0) - F_X(t_0) \prod_{j \neq i} F_{X_j}(t_0)}{2} < \prod_{j=1}^n F_{X_j}(t_0).$$

Taking $\delta = \frac{\prod_{j=1}^n F_{X_j}(t_0) - F_X(t_0) \prod_{j \neq i} F_{X_j}(t_0)}{2} > 0$, then:

$$\prod_{j=1}^n F_{X_j}(t) - F_X(t) \prod_{j \neq i} F_{X_j}(t) \geq \delta > 0$$

for any $t \in [t_0, t_0 + \varepsilon]$. Moreover, since F_{X_i} is strictly increasing in $[a, b]$ it is also strictly increasing in $[t_0, t_0 + \varepsilon]$, and therefore $P(X_i \in [t_0, t_0 + \varepsilon]) > 0$. ∎

Proposition 3.93 *Let $X, X_1, \ldots, X_n$ be $n+1$ independent real-valued random variables satisfying Equations* (3.25) *and* (3.26). *Then, if $X \succ_{FSD} X_i$ for any $i = 1, \ldots, n$ it holds that $X \succ_{SP} [X_1, \ldots, X_n]$.*

Proof Since $X \succ_{FSD} X_i$ implies $X \succeq_{FSD} X_i$, we know that $X \succeq_{SP} [X_1,\dots,X_n]$. Taking into account the previous result, it suffices to prove that $E[F_X(X_i)\prod_{j\neq i} F_{X_j}(X_i)] < E[\prod_{j=1}^{n} F_{X_j}(X_i)]$ for $i = 1,\dots,n$, since this implies that:

$$Q_n(X,[X_1,\dots,X_n]) \geq Q_n(X_i,[X,X_{-i}]) \text{ for } i = 1,\dots,n.$$

Using the previous lemma, we can assume there is t_0 such that $F_X(t_0) < F_{X_i}(t_0)$ and $F_{X_j}(t_0) > 0$ for any $j \neq i$.

Consider two cases:

- Assume that $P(X_i = t_0) > 0$. Then:

$$\begin{aligned}
E\Big[F_X(X_i)\prod_{j\neq i}F_{X_j}(X_i)\Big] &= \int F_X(X_i)\prod_{j\neq i}F_{X_j}\,dF_{X_i}\\
&= \int_{\mathbb{R}\setminus\{t_0\}} F_X(X_i)\prod_{j\neq i}F_{X_j}\,dF_{X_i} + \int_{\{t_0\}} F_X(X_i)\prod_{j\neq i}F_{X_j}\,dF_{X_i}\\
&\leq \int_{\mathbb{R}\setminus\{t_0\}} \prod_{j=1}^{n}F_{X_j}\,dF_{X_i} + P(X_i = t_0)F_X(X_i)(t_0)\prod_{j\neq i}F_{X_j}(t_0)\\
&< \int_{\mathbb{R}\setminus\{t_0\}} \prod_{j=1}^{n}F_{X_j}\,dF_{X_i} + P(X_i = t_0)\prod_{j=1}^{n}F_{X_j}(t_0)\\
&= \int_{\mathbb{R}\setminus\{t_0\}} \prod_{j=1}^{n}F_{X_j}\,dF_{X_i} + \int_{\{t_0\}} \prod_{j=1}^{n}F_{X_j}\,dF_{X_i}\\
&= E\Big[\prod_{j=1}^{n}F_{X_j}(X_i)\Big].
\end{aligned}$$

- Assume now that there is not t_0 satisfying the conditions and such that $P(X_i = t_0) = 0$. By the previous lemma, there is an interval $[a,b]$ and $\varepsilon > 0$ such that

$$\prod_{j=1}^{n}F_{X_j}(t) - F_X(t)\prod_{j\neq i}F_{X_j}(t) \geq \varepsilon > 0$$

for any $t \in [a, b]$ and $P(X_i \in [a, b]) > 0$. Then:

$$
\begin{aligned}
E\left[F_X(X_i)\prod_{j\neq i} F_{X_j}(X_i)\right] &= \int F_X(X_i)\prod_{j\neq i} F_{X_j}\,dF_{X_i}\\
&= \int_{\mathbb{R}\setminus[a,b]} F_X(X_i)\prod_{j\neq i} F_{X_j}\,dF_{X_i} + \int_{[a,b]} F_X(X_i)\prod_{j\neq i} F_{X_j}\,dF_{X_i}\\
&\leq \int_{\mathbb{R}\setminus[a,b]} \prod_{j=1}^{n} F_{X_j}\,dF_{X_i} + \int_{[a,b]} \left(\prod_{j=1}^{n} F_{X_j} - \varepsilon\right) dF_{X_i}\\
&= \int \prod_{j=1}^{n} F_{X_j}\,dF_{X_i} + \varepsilon P(X_i \in [a, b]) < E\left[\prod_{j=1}^{n} F_{X_j}(X_i)\right].
\end{aligned}
$$

We have seen that $X \succ_{FSD} X_i$ for any $i = 1, \ldots, n$, implies that $X \succ_{SP} [X_1, \ldots, X_n]$ when the random variables are independent. Since general statistical preference is based on the joint distribution, and as a consequence takes into account the possible stochastic dependencies between the variables, we are going to study a number of cases where the variables are not independent. In the remainder of this subsection we shall focus on comonotonic random variables.

In Equation (3.6) of Proposition 3.16 we saw that the probabilistic relation $Q(X, Y)$ for two continuous and comonotonic random variables is given by:

$$Q(X, Y) = \int_{x : F_X(x) < F_Y(x)} f_X(x)\,dx + \frac{1}{2}\int_{x : F_X(x) = F_Y(x)} f_X(x)\,dx,$$

where f_X denotes the density function of X.

In a similar manner, we can extend this expression to the functional $Q_n(\cdot, [\cdot])$. In order to do this, we must first introduce the notion of Dirac-delta functional. Let us consider the function $H_a : \mathbb{R} \to [0, 1]$ given by:

$$H_a(x) = \begin{cases} 0 & \text{if } x < a. \\ 1 & \text{if } x \geq a. \end{cases}$$

The Dirac-delta functional δ_a (see [66]) associated to H_a is an application that satisfies:

- $\delta_a(t) = 0$ for every $t \neq a$ and
- $\int_{\mathbb{R}} \delta_a(t)\,dt = 1$.

In such a case, it holds that:

$$H_a(x) = \int_{-\infty}^{x} \delta_a(t)\,d(t) \quad \text{for every } x \in \mathbb{R}. \tag{3.27}$$

This functional is not a real-valued function because it does not take a real value in a. It plays the role of the density function for a probability distribution that takes the value a with probability 1, and we shall use it in the proof of the following lemma.

Lemma 3.94 *Let $X, X_1, \ldots, X_n$ be absolutely continuous and comonotonic real-valued random variables satisfying Equation* (3.25). *Then*

$$Q_n(X, [X_1, \ldots, X_n]) = \int_{x: F_X(x) < F_{X_1}(x), \ldots, F_{X_n}(x)} f_X(x) \mathrm{d}(x).$$

Proof By Corollary 3.89, it holds that:

$$Q_n(X, [X_1, \ldots, X_n]) = P(X > \max(X_1, \ldots, X_n)).$$

Since the random variables are comonotonic, their joint distribution function F is given by:

$$F(x, x_1, \ldots, x_n) = \min(F_X(x), F_{X_1}(x_1), \ldots, F_{X_n}(x_n))$$

for every $x, x_1, \ldots, x_n \in \mathbb{R}$. Let us compute the distribution function of $\max(X_1, \ldots, X_n)$ and X, denoted by F:

$$\begin{aligned} F(x, y) &= P(X \leq x, \max(X_1, \ldots, X_n) \leq y) \\ &= P(X \leq x, X_1 \leq y, \ldots, X_n \leq y) = F(x, y, \ldots, y). \end{aligned}$$

Thus, this distribution function can be expressed by:

$$\begin{aligned} F(x, y) &= F(x, y, \ldots, y) = \min(F_X(x), F_{X_1}(y), \ldots, F_{X_n}(y)) \\ &= \begin{cases} F_X(x) & \text{if } F_X(x) \leq \min(F_{X_1}(y), \ldots, F_{X_n}(y)). \\ \min(F_{X_1}(y), \ldots, F_{X_n}(y)) & \text{if } F_X(x) > \min(F_{X_1}(y), \ldots, F_{X_n}(y)). \end{cases} \end{aligned}$$

Equivalently,

$$F(x, y) = \begin{cases} F_X(x) & \text{if } y \geq h^{-1]}(F_X(x)), \\ \min(F_{X_1}(y), \ldots, F_{X_n}(y)) & \text{if } y < h^{-1]}(F_X(x)), \end{cases}$$

where $h^{-1]}$ denotes the pseudo-inverse of the function h given by:

$$h(y) = \min(F_{X_1}(y), \ldots, F_{X_n}(y)) \text{ for every } y \in \mathbb{R}.$$

Note that the pseudo-inverse is well-defined since h is an increasing function. Now, $\frac{\partial F}{\partial x}(x, y) = 0$ for every (x, y) satisfying $y < h^{-1]}(F_X(x))$. Moreover, if we restrict to the points (x, y) such that $y \geq h^{-1]}(F_X(x))$, we obtain that:

$$\frac{\partial F}{\partial x}(x, y) = f_X(x).$$

Thus, if we assume that:

$$\frac{\partial F}{\partial x}(x,y) = \begin{cases} 0 & \text{if } y < h^{-1}(F_X(x)), \\ f_X(x) & \text{if } y \geq h^{-1}(F_X(x)), \end{cases}$$

then:

$$\frac{\partial^2 F}{\partial x \partial y}(x,y) = f_X(x)\delta\left(y - h^{-1}(F_X(x))\right).$$

As this distribution plays the role of the density function of $\max(X_1,\ldots,X_n)$ and X, using Equation (3.27) we can compute the value of $Q_n(X,[X_1,\ldots,X_n])$:

$$\begin{aligned}
Q_n(X,[X_1,\ldots,X_n]) &= P\left(X > \max(X_1,\ldots,X_n)\right) \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} f_X(x)\delta\left(y - h^{-1}(F_X(x))\right) I_{x>y}(y)\,dy\,dx \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} f_X(x)\delta\left(y - h^{-1}(F_X(x))\right) \lim_n I_{\{x-y\geq 1/n\}}(y)\,dy\,dx \\
&= \int_{\mathbb{R}} \lim_n \int_{-\infty}^{x-1/n} f_X(x)\delta\left(y - h^{-1}(F_X(x))\right) dy\,dx \\
&= \int_{\mathbb{R}} \lim_n f_X(x) I_{\{x-1/n \geq h^{-1}(F_X(x))\}}(x)\,dy\,dx \\
&= \int_{\mathbb{R}} f_X(x) I_{\{x>h^{-1}(F_X(x))\}}(x)\,dy\,dx \\
&= \int_{\{F_X(x)<F_{X_1}(x),\ldots,F_{X_n}(x)\}} f_X(x)\,dx,
\end{aligned}$$

where the last equality holds applying the Theorem of Monotone Convergence. ∎

Theorem 3.95 *Let $X, X_1,\ldots,X_n$ be $n+1$ absolutely continuous and comonotonic real-valued random variables satisfying Equations* (3.25) *and* (3.26). *If $X \succ_{FSD} X_i$ for $i = 1,\ldots,n$, then $X \succ_{SP} [X_1,\ldots,X_n]$. Moreover, in that case $Q_n(X,[X_1,\ldots,X_n]) = 1$.*

Proof Since $X \succ_{FSD} X_i$ for every $i = 1,\ldots,n$, then $F_X(x) \leq F_{X_i}(x)$ for every $x \in \mathbb{R}$ and $i = 1,\ldots,n$. Applying the previous lemma we obtain that:

$$Q_n(X_i,[X,X_{-i}]) = \int_{\{F_{X_i}(x)<F_X(x),F_{X_j}(x):j\neq i\}} f_{X_i}(x)\,dx = 0.$$

Thus, $Q_n(X,[X_1,\ldots,X_n]) > Q_n(X_i,[X,X_{-i}]) = 0$ for every $i = 1,\ldots,n$. Since

$$Q_n(X,[X_1,\ldots,X_n]) + \sum_{i=1}^{n} Q_n(X_i,[X,X_{-i}]) = 1,$$

it holds that:

$$Q_n(X, [X_1, \ldots, X_n]) = 1.$$

Then, X SP $[X_1, \ldots, X_n]$. ∎

Let us now investigate the case in which the random variables $X, X_1, \ldots, X_n$ are comonotonic and discrete with finite supports. When $n=1$, De Meyer et al. proved (see Proposition 3.20) that the supports of the variables can be expressed by $S_X = \{x_1, \ldots, x_m\}$ and $S_{X_1} = \{x_1^{(1)}, \ldots, x_m^{(1)}\}$ such that

$$P(X = x_i, X_1 = x_i^{(1)}) = P(X = x_i) = P(X_1 = x_i^{(1)}) \text{ for any } i = 1, \ldots, m.$$

We are going to prove the a similar expression can be found when $n \geq 2$.

Lemma 3.96 *Let $X, X_1, \ldots, X_n$ be $n+1$ discrete and countermonotonic real-valued random variables with finite supports. Then, their supports can be expressed by*

$$S_X = \{x_1, \ldots, x_m\},\ S_{X_1} = \{x_1^{(1)}, \ldots, x_m^{(1)}\},\ \ldots,\ S_{X_n} = \{x_1^{(n)}, \ldots, x_m^{(n)}\}, \tag{3.28}$$

and

$$P(X = x_i, X_1 = x_i^{(1)}, \ldots, X_n = x_i^{(n)}) = P(X = x_i) = \ldots = P(X_n = x_i^{(n)}), \tag{3.29}$$

for any $i = 1, \ldots, n$.

Proof We apply induction on n. First of all, when $n = 1$, this lemma becomes Proposition 3.20. Assume then that the result holds for $n-1$. Consider the variables $X, X_1, \ldots, X_n$. Apply the induction hypothesis on $X, X_1, \ldots, X_{n-1}$. Then, the supports of these variables can be expressed as in Equation (3.28), and they also satisfy Equation (3.29). Now, apply Proposition 3.20 to X (with the new support) and X_n. Then, if in this process we duplicate an element x_i, we also duplicate the elements $x_i^{(j)}$ for any $j = 1, \ldots, n-1$, and we adapt the probabilities in order to obtain the equalities:

$$P(X = x_i) = P(X_n = x_i^{(1)}) = \ldots = P(X_n = x_i^{(n)}).$$

Finally, let us prove that

$$P(X = x_i, X_1 = x_i^{(1)}, \ldots, X_n = x_i^{(n)}) = P(X = x_i).$$

For this aim, note that

$$F_X(x_j) = P(X = x_1) + \ldots + P(X = x_j) = P(X_i = x_1^{(i)}) + \ldots + P(X_i = x_j^{(i)}) = F_{X_i}(x_j^{(i)})$$

for any $j = 1, \ldots, m$ and $i = 1, \ldots, n$. Then:

$$\begin{aligned} F_{X, X_1, \ldots, X_n}(x_{i_0}, x_{i_1}^{(1)}, \ldots, x_{i_n}^{(n)}) &= \min(F_X(x_{i_0}), F_{X_1}(x_{i_1}^{(1)}), \ldots, F_{X_n}(x_{i_n}^{(n)})) \\ &= \min(F_X(x_{i_0}), F_X(x_{i_1}^{(1)}), \ldots, F_X(x_{i_n}^{(n)})) \\ &= F_X(\min_{k=0,\ldots,n}(x_{i_k})). \end{aligned}$$

In particular, when $i_0 = i_1 = \ldots, i_n$, the previous expression becomes:

$$F_{X,X_1,\ldots,X_n}(x_i, x_i^{(1)}, \ldots, x_i^{(n)}) = F_X(x_i).$$

Now, consider $(x_{i_0}, x_{i_1}^{(1)}, \ldots, x_{i_n}^{(n)})$, and assume that there are k, l such that $i_k = i_l$. Since in the proof of Proposition 3.20 (see [54, Proposition 2]) it is showed that $P(X_k = x_{i_k}^{(k)}, X_l = x_{i_l}^{(l)}) = 0$, we deduce that:

$$P(X = x_{i_0}, X_1 = x_{i_1}^{(1)}, \ldots, X_n = x_{i_n}^{(n)}) \leq P(X_k = x_{i_k}^{(k)}, X_l = x_{i_l}^{(l)}) = 0.$$

Consequently:

$$\begin{aligned} P(X = x_i, X_i = x_i^{(1)}, \ldots, X_i = x_i^{(n)}) &= F(x_i, x_i^{(1)}, \ldots, x_i^{(n)}) - F(x_{i-1}, x_{i-1}^{(1)}, \ldots, x_{i-1}^{(n)}). \\ &= F_X(x_i) - F_X(x_{i-1}) = P(X = x_i). \end{aligned}$$

Next result gives an expression of the probabilistic relation, generalizing Equation (3.8).

Proposition 3.97 *Consider $n+1$ discrete and comonotonic real-valued random variables $X, X_1, \ldots, X_n$ with finite supports. Applying the previous lemma, we can assume that the supports are expressed as in Equation (3.28) satisfying Equation (3.29). Then:*

$$Q_n(X, [X_1, \ldots, X_n]) = \sum_{i=1}^{n} P(X = x_i)\delta_i,$$

where

$$\delta_i = \begin{cases} 0, & \text{if } x_i > x_i^{(1)}, \ldots, x_i^{(n)}. \\ \frac{1}{2}, & \text{if } x_i = x_i^{(j)} > x_i^{(k)}, \text{ for any } k \neq j. \\ \frac{1}{3}, & \text{if } x_i = x_i^{(j_1)} = x_i^{(j_2)} > x_i^{(k)}, \text{ for any } k \neq j_1, j_2. \\ \ldots & \\ \frac{1}{n}, & \text{if } x_i = x_i^{(1)} = \ldots = x_i^{(n)}. \end{cases}$$

Proof Taking into account Equation (3.29), it holds that:

$$\begin{aligned} P(X > X_1, \ldots, X_n) &= \sum_{i_0=1}^{m} \cdots \sum_{i_n=1}^{m} P(X = x_{i_0}, X_1 = x_{i_1}^{(1)}, \ldots, X_n = x_{i_n}^{(n)}) I_{x_{i_0} > x_{i_1}^{(1)}, \ldots, x_{i_n}^{(n)}} \\ &= \sum_{i=1}^{m} P(X = x_i, X_1 = x_i^{(1)}, \ldots, X_n = x_i^{(n)}) I_{x_i > x_i^{(1)}, \ldots, x_i^{(n)}}. \end{aligned}$$

Similarly:

$$
\begin{aligned}
P(X = X_{l_1} = \ldots = X_{l_k} > X_j : j = l_1, \ldots, l_k) &= \sum_{i_0=1}^{m} \cdots \sum_{i_n=1}^{m} P(X = x_{i_0}, X_1 = x^{(1)}_{i_1}, \ldots, X_n = x^{(n)}_{i_n}) I_A \\
&= \sum_{i=1}^{m} P(X = x_i, X_1 = x^{(1)}_i, \ldots, X_n = x^{(n)}_i) I_B,
\end{aligned}
$$

where A and B are defined by:

$$
\begin{aligned}
A &= \{x_{i_0} = x^{(l_1)}_{i_{l_1}} = \ldots = x^{(l_k)}_{i_{l_k}} > x^{(j)}_{i_j} : j = l_1, \ldots, k\} \text{ and} \\
B &= \{x_i = x^{(l_1)}_i = \ldots = x^{(l_k)}_i > x^{(j)}_i : j = l_1, \ldots, k\}.
\end{aligned}
$$

Then:

$$
\begin{aligned}
Q_n(X, [X_1, \ldots, X_n]) &= \sum_{\substack{k=0,\ldots,n \\ 1 \leq i_1 < \ldots < i_k \leq n}} \frac{1}{k+1} P(X = X_{i_1} = \ldots = X_{i_k} > \max_{j = i_1,\ldots,i_k}(X_j)) \\
&= \sum_{i=1}^{m} \sum_{i=1}^{n} P(X = x_i)\delta_i.
\end{aligned}
$$

Remark 3.98 *In this result we have not imposed Equations* (3.25) *and* (3.26)*, and thus, it is applicable for all discrete comonotonic random variables with finite supports.*

Using this lemma, we can prove that when the random variables are comonotonic and discrete with finite supports, first degree stochastic dominance also implies general statistical preference.

Theorem 3.99 *Let $X, X_1, \ldots, X_n$ be $n+1$ discrete comonotonic real-valued random variables with finite supports. Then $X \ _{\mathrm{FSD}}\ X_i$ for $i = 1, \ldots, n$ implies $X \ _{\mathrm{SP}}$ $[X_1, \ldots, X_n]$.*

Proof Using the previous lemma, the supports of $X, X_1, \ldots, X_n$ can be expressed as in Equation (3.28) satisfying Equation (3.29). If $X \ _{\mathrm{FSD}}\ X_i$, we have seen in the proof of Theorem 3.51 that $\{i : x_i < x^{(j)}_i\} =$ for $j = 1, \ldots, n$. Using the previous proposition:

$$
\begin{aligned}
Q_n(X_i, [X, X_{-i}]) &= \sum_{\substack{k=0,\ldots,n-1 \\ 1 \leq i_1 < \ldots < i_k \leq n}} \frac{1}{k+1} P(X_i = X = X_{i_1} = \ldots = X_{i_k} > \max_{j = i, i_1,\ldots,i_{k-1}}(X_j)) \\
&\leq \sum_{\substack{k=0,\ldots,n \\ 1 \leq i_1 < \ldots < i_k \leq n}} \frac{1}{k+1} P(X = X_{i_1} = \ldots = X_{i_k} > \max_{j = i_1,\ldots,i_k}(X_j)) = Q(X, Y),
\end{aligned}
$$

and this for any $i = 1, \dots, n$. Then, $X \succeq_{SP} [X_1, \dots, X_n]$. ∎

Finally, let us prove that when X is strictly preferred to any X_i with respect to first degree stochastic dominance, it is also preferred to $[X_1, \dots, X_n]$ with respect to the general statistical preference.

Proposition 3.100 *Let $X, X_1, \dots, X_n$ be $n+1$ discrete comonotonic real-valued random variables with finite supports. Then $X \succ_{FSD} X_i$ for $i = 1, \dots, n$ implies $X \succ_{SP} [X_1, \dots, X_n]$.*

Proof Let us prove that $Q_n(X, [X_1, \dots, X_n]) > Q_n(X_i, [X, X_{-i}])$ for $i = 1, \dots, n$. From the proof of the previous result, it suffices to prove that there are k and l such that

$$x_k = x_k^{(j_1)} = \dots = x_k^{(j_l)} > x_k^{(i)}, x_k^{(j)}, \text{ such that } j \neq i, j_1, \dots, j_l.$$

Since $X \succ_{FSD} X_i$, there is $x_k^{(i)}$ such that $F_X(x_k^{(i)}) < F_{X_i}(x_k^{(i)})$. Furthermore:

$$F_{X_i}(x_k^{(i)}) = P(X_i = x_1^{(i)}) + \dots + P(X_i = x_k^{(i)}) = P(X = x_1) + \dots + P(X = x_k) = F_X(x_k).$$

Then, $x_k > x_k^{(i)}$. Then, there is l such that

$$x_k = x_k^{(j_1)} = \dots = x_k^{(j_l)} > x_k^{(i)}, x_k^{(j)}, \text{ such that } j \neq i, j_1, \dots, j_l,$$

and this proves that $Q_n(X, [X_1, \dots, X_n]) > Q_n(X_i, [X, X_{-i}])$, for $i = 1, \dots, n$. Hence $X \succ_{SP} [X_1, \dots, X_n]$. ∎

Random variables coupled by Archimedean copulas

Consider $n+1$ absolutely continuous random variables $X, X_1, \dots, X_n$ coupled by an Archimedean copula C with generator ϕ. In that case, Equation (2.9) implies that the joint distribution function, F, is given by:

$$F(x, x_1, \dots, x_n) = \phi^{[-1]}\left(\phi(F_X(x)) + \phi(F_{X_1}(x_1)) + \dots + \phi(F_{X_n}(x_n))\right).$$

Let us try to differentiate this function.

$$\frac{\partial F}{\partial x}(x, x_1, \dots, x_n) =$$
$$\phi^{[-1]\prime}\left(\phi(F_X(x)) + \phi(F_{X_1}(x_1)) + \dots + \phi(F_{X_n}(x_n))\right)\phi'(F_X(x))f_X(x).$$

Note that $\phi^{[-1]}(t)$ equals $\phi^{-1}(t)$ whenever $t \in [0, \phi(0))$ and $\phi^{[-1]}(t) = 0$ otherwise. If we continue differentiating with respect to $x_1, \dots, x_n$, we obtain the following expression:

$$\frac{\partial^2 F}{\partial x \partial x_1}(x, x_1, \ldots, x_n) = \phi^{-1]} \left(\phi(F_X(x)) + \phi(F_{X_1}(x_1)) + \ldots + \phi(F_{X_n}(x_n))\right) \phi(F_X(x))\phi(F_{X_1}(x_1)) f_X(x) f_{X_1}(x_1).$$

$$\ldots$$

$$\frac{\partial^{n+1} F}{\partial x \partial x_1 \ldots \partial x_n}(x, x_1, \ldots, x_n) = \phi^{-1](n+1)} \left(\phi(F_X(x)) + \phi(F_{X_1}(x_1)) + \ldots + \phi(F_{X_n}(x_n))\right) \phi(F_X(x)) \prod_{i=1}^{n} \phi(F_{X_i}(x_i)) f_{X_i}(x_i) f_X(x).$$

Thus, function $f(x, x_1, \ldots, x_n) = \frac{\partial^{n+1} F}{\partial x \partial x_1 \ldots \partial x_n}(x, x_1, \ldots, x_n)$ is the density function of $X, X_1, \ldots, X_n$ whenever $f = 0$, since it is the $n+1$ derivative of F, and the $n+1$ integral over $\mathbb{R}^{n+1}$ equals 1. In addition, f becomes the density function of Equation (3.10). Note that $f = 0$ when $\phi^{-1](n+1)}(t) > 0$ for some $t \in \mathbb{R}$. Moreover, if f is the joint density, $P(X = X_i) = P(X_i = X_j) = 0$ for every i, j $(i = j)$. Consequently, for such variables it holds that:

$$\begin{aligned} Q_n(X, [X_1, \ldots, X_n]) &= P(X > \max(X_1, \ldots, X_n)) \\ &= P(X \geq \max(X_1, \ldots, X_n)) = Q(X, \max(X_1, \ldots, X_n)). \end{aligned}$$

Using the joint density function f, we can prove the following result.

Theorem 3.101 *Let $X, X_1, \ldots, X_n$ be $n+1$ absolutely continuous random variables coupled by an Archimedean copula C generated by ϕ, that satisfies $\phi^{-1](n+1)} = 0$. Then, if $X \;{}_{FSD}\; X_i$ for every $i = 1, \ldots, n$, then $X \;{}_{SP}\; [X_1, \ldots, X_n]$.*

Proof We know that $X \;{}_{SP}\; [X_1, \ldots, X_n]$ if and only if

$$P(X \geq \max(X_1, \ldots, X_n)) \geq P(X_i \geq \max(X, X_{-i})),$$

for every $i = 1, \ldots, n$. Let us compute $P(X \geq \max(X_1, \ldots, X_n))$.

$$\begin{aligned} P(X \geq \max(X_1, \ldots, X_n)) &= \int_{\mathbb{R}} \int_{-\infty}^{x} \cdots \int_{-\infty}^{x} f(x, x_1, \ldots, x_n)\, dx_n \ldots dx_1\, dx \\ &= \int_{\mathbb{R}} \phi^{-1]} \left(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x))\right) \phi(F_X(x)) f_X(x)\, dx. \end{aligned}$$

If we consider

$$u = \phi^{-1]} \left(\phi(F_X(x)) + \phi(F_{X_1}(x)) + \ldots + \phi(F_{X_n}(x))\right),$$
$$dv = \phi(F_X(x)) f_X(x)\, dx,$$

and we make a change of variable, we obtain the following expression:

$$\begin{aligned} &P(X \geq \max(X_1, \ldots, X_n)) = \\ &\quad 1 - \int_{\mathbb{R}} \phi^{-1]} \left(\phi(F_X(x)) + \phi(F_{X_1}(x)) + \ldots + \phi(F_{X_n}(x))\right) \phi(F_X(x)) \\ &\quad \left(\phi(F_X(x)) f_X(x) + \sum_{i=1}^{n} \phi(F_{X_i}(x)) f_{X_i}(x)\right) dx. \end{aligned}$$

Now, since $X \succeq_{FSD} X_i$, then $F_X \leq F_{X_i}$, and consequently, as $\phi(F_X(x)) \geq \phi(F_{X_i}(x))$ (ϕ is decreasing), ϕ' is negative and $\phi^{-1\prime}$ is positive, it holds that:

$$\begin{aligned}
&P(X \geq \max(X_1,\dots,X_n)) \geq \\
&\quad 1-\int_{\mathbb{R}} \phi^{[-1]\prime}\Big(\phi(F_X(x))+\phi(F_{X_1}(x))+\dots+\phi(F_{X_n}(x))\Big)\ \phi(F_{X_i}(x)) \\
&\qquad \Big(\phi'(F_X(x))f_X(x)+\sum_{i=1}^{n}\phi'(F_{X_i}(x))f_{X_i}(x)\Big)\,dx.
\end{aligned}$$

Following the same lines we can also find the expression of $P(X_i \geq \max(X, X_{-i}))$:

$$\begin{aligned}
&P(X_i \geq \max(X, X_{-i}))= \\
&\quad 1-\int_{\mathbb{R}} \phi^{[-1]\prime}\Big(\phi(F_X(x))+\phi(F_{X_1}(x))+\dots+\phi(F_{X_n}(x))\Big)\ \phi(F_{X_i}(x)) \\
&\qquad \Big(\phi'(F_X(x))f_X(x)+\sum_{i=1}^{n}\phi'(F_{X_i}(x))f_{X_i}(x)\Big)\,dx.
\end{aligned}$$

We conclude that:

$$P(X \geq \max(X_1,\dots,X_n)) \geq P(X_i \geq \max(X, X_{-i})),$$

and consequently $X \succeq_{SP} [X_1,\dots,X_n]$. ∎

Finally, let us see that when the Archimedean copula is strict, strict statistical first degree stochastic dominance also implies strict statistical preference.

Proposition 3.102 *Let $X, X_1, \dots, X_n$ be $n+1$ absolutely continuous random variables coupled by an strict Archimedean copula C generated by ϕ, that satisfies $\phi^{-1]\,(n+1)} = 0$. Then, if $X \succ_{FSD} X_i$ for every $i = 1,\dots,n$, then $X \succ_{SP} [X_1,\dots,X_n]$.*

Proof By Lemma 3.48, since $X \succ_{FSD} X_i$, there is an interval $[a, b]$ such that $F_X(t) < F_{X_i}(t)$ for any $t \in [a, b]$ and $P(X_i \in [a, b]) > 0$. Furthermore, we can assume that F_{X_i} is strictly increasing in such interval (otherwise it suffices to consider a subinterval of $[a, b]$ where this function is strictly increasing).

We have seen in the previous proof that

$$\begin{aligned}
&P(X \geq \max(X_1,\dots,X_n))= \\
&\quad 1-\int_{\mathbb{R}} \phi^{[-1]\prime}\Big(\phi(F_X(x))+\phi(F_{X_1}(x))+\dots+\phi(F_{X_n}(x))\Big)\ \phi(F_X(x)) \\
&\qquad \Big(\phi'(F_X(x))f_X(x)+\sum_{i=1}^{n}\phi'(F_{X_i}(x))f_{X_i}(x)\Big)\,dx
\end{aligned}$$

and

$$P(X_i \geq \max(X, X_{-i})) =$$
$$1 - \int_{\mathbb{R}} \phi^{[-1]}\Big(\phi(F_X(x)) + \phi(F_{X_1}(x)) + \ldots + \phi(F_{X_n}(x)) \;\; \phi(F_{X_i}(x))\Big)$$
$$\Big(\phi(F_X(x)) f_X(x) + \sum_{i=1}^{n} \phi(F_{X_i}(x)) f_{X_i}(x)\Big) \mathrm{d}x.$$

Then, in order to prove that $Q_n(X, [X_1, \ldots, X_n]) > Q_n(X_i, [X, X_{-i}])$, it suffices to prove that:

$$1 - \int_{\mathbb{R}} \phi^{[-1]}\Big(\phi(F_X(x)) + \phi(F_{X_1}(x)) + \ldots + \phi(F_{X_n}(x)) \;\; \phi(F_X(x))\Big)$$
$$\Big(\phi(F_X(x)) f_X(x) + \sum_{i=1}^{n} \phi(F_{X_i}(x)) f_{X_i}(x)\Big) \mathrm{d}x$$
$$> 1 - \int_{\mathbb{R}} \phi^{[-1]}\Big(\phi(F_X(x)) + \phi(F_{X_1}(x)) + \ldots + \phi(F_{X_n}(x)) \;\; \phi(F_{X_i}(x))\Big)$$
$$\Big(\phi(F_X(x)) f_X(x) + \sum_{i=1}^{n} \phi(F_{X_i}(x)) f_{X_i}(x)\Big) \mathrm{d}x,$$

or equivalently:

$$\int_{\mathbb{R}} \phi^{[-1]}\Big(\phi(F_X(x)) + \phi(F_{X_1}(x)) + \ldots + \phi(F_{X_n}(x)) \;\; \phi(F_X(x))\Big)$$
$$\Big(\phi(F_X(x)) f_X(x) + \sum_{i=1}^{n} \phi(F_{X_i}(x)) f_{X_i}(x)\Big) \mathrm{d}x$$
$$< \int_{\mathbb{R}} \phi^{[-1]}\Big(\phi(F_X(x)) + \phi(F_{X_1}(x)) + \ldots + \phi(F_{X_n}(x)) \;\; \phi(F_{X_i}(x))\Big)$$
$$\Big(\phi(F_X(x)) f_X(x) + \sum_{i=1}^{n} \phi(F_{X_i}(x)) f_{X_i}(x)\Big) \mathrm{d}x.$$

By the proof of the previous theorem, we know that:

$$\int_{\mathbb{R}} \phi^{[-1]}\Big(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \;\; \phi(F_X(x))\Big) \phi(F_X(x)) f_X(x) \mathrm{d}x \leq$$
$$\int_{\mathbb{R}} \phi^{[-1]}\Big(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \;\; \phi(F_X(x))\Big) \phi(F_{X_i}(x)) f_{X_i}(x) \mathrm{d}x$$

and

$$\int_{\mathbb{R}} \phi^{[-1]}\Big(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \;\; \phi(F_{X_j}(x))\Big) \phi(F_X(x)) f_{X_j}(x) \mathrm{d}x \leq$$
$$\int_{\mathbb{R}} \phi^{[-1]}\Big(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \;\; \phi(F_{X_j}(x))\Big) \phi(F_{X_i}(x)) f_{X_j}(x) \mathrm{d}x.$$

Now, let us see that for $j = i$, the previous inequality is strict. For any $t \in [a, b]$

$$
\begin{aligned}
F_{X_i}(t) < F_X(t) &\overset{\phi \text{ decr.}}{\Rightarrow} \phi(F_{X_i}(t)) > \phi(F_X(t)) \\
&\overset{\phi' < 0}{\Rightarrow} \phi'(F_{X_i}(t))\phi(F_{X_i}(t)) < \phi'(F_{X_n}(t))\phi(F_X(t)) \\
&\overset{(\phi^{-1})' < 0}{\Rightarrow} \phi^{-1]}\left(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \phi'(F_{X_i}(t))\phi(F_{X_i}(t))\right) > \\
&\quad \phi^{-1]}\left(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \phi'(F_X(t))\phi(F_X(t))\right).
\end{aligned}
$$

Then, there is $\varepsilon > 0$ and $[a_1, b_1] \subseteq [a, b]$ such that

$$
\begin{aligned}
&\phi^{-1]}\left(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \phi'(F_{X_i}(t))\phi(F_{X_i}(t))\right) - \\
&\qquad \phi^{-1]}\left(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \phi'(F_X(t))\phi(F_X(t))\right) \geq \varepsilon > 0
\end{aligned}
$$

for any $t \in [a_1, b_1]$. Then:

$$
\begin{aligned}
&\int_{\mathbb{R}} \phi^{-1]}\left(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \phi(F_{X_i}(x))\right) \phi'(F_{X_i}(x)) f_{X_i}(x)\,dx = \\
&\int_{\mathbb{R}\setminus[a_1,b_1]} \phi^{-1]}\left(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \phi(F_{X_i}(x))\right) \phi'(F_{X_i}(x)) f_{X_i}(x)\,dx \\
&+ \int_{[a_1,b_1]} \phi^{-1]}\left(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \phi(F_{X_i}(x))\right) \phi'(F_{X_i}(x)) f_{X_i}(x)\,dx \\
&\geq \int_{\mathbb{R}} \phi^{-1]}\left(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \phi(F_{X_i}(x))\right) \phi'(F_X(x)) f_{X_i}(x)\,dx + \\
&\int_{[a_1,b_1]} \varepsilon f_{X_i}(x)\,dx = \\
&\int_{\mathbb{R}} \phi^{-1]}\left(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \phi(F_{X_i}(x))\right) \phi'(F_X(x)) f_{X_i}(x)\,dx + \\
&\varepsilon P(X_i \in [a_1, b_1]) > \\
&\int_{\mathbb{R}} \phi^{-1]}\left(\phi(F_X(x)) + \sum_{k=1}^{n} \phi(F_{X_k}(x)) \phi(F_{X_i}(x))\right) \phi'(F_X(x)) f_{X_i}(x)\,dx.
\end{aligned}
$$

Therefore, $Q_n(X, [X_1, \ldots, X_n]) > Q_n(X_i, [X, X_{-i}])$, and then we can conclude that $X \succ_{SP} [X_1, \ldots, X_n]$. ■

We have seen several situations where $X \succeq_{FSD} X_i$, $i = 1, \ldots, n$ implies $X \succeq_{SP} [X_1, \ldots, X_n]$. However, this implication does not hold in general, as we can see in the following example.

Example 3.103 *We have seen in Example3.43 two random variables X and Y such that X $_{FSD}$ Y and Y $_{SP}$ X. These random variableswere defined by:*

X/Y	0	1	2
0	0.2	0.15	0
1	0	0.2	0.15
2	0.2	0	0.1

It holds that $Q(X, Y) = 0.45$. Let us modify this example to show that if there isa random variable X that stochastical ly dominates any other random variables, it may not be the preferred with respect to the general statistical preference. Consider $X_1, \ldots, X_n$ equal ly distributed such that they take a fixed value $c<0$ with probability 1. Since X and Y greater than $X_1, \ldots, X_n$ with probabilityone, X $_{FSD}$ X_i for $i = 1, \ldots, n$, and it holds that:

$$\begin{aligned} Q_{n+1}(X, [Y, X_1, \ldots, X_n]) &= P(X > \max(Y, X_1, \ldots, X_n)) \\ &+ \tfrac{1}{2} P(X = Y > \max(X_1, \ldots, X_n)) \\ &= P(X > Y) + \tfrac{1}{2} P(X = Y) = Q(X, Y) = 0.45. \end{aligned}$$

Similarly, $Q_{n+1}(Y, [X, X_1, \ldots, X_n]) = Q(Y, X) = 0.55$. Therefore, X $_{FSD}$ Y, X $_{FSD}$ X_i for $i = 1, \ldots, n$ but X $_{SP}$ $[Y, X_1, \ldots, X_n]$.

To conclude this section we are going to se e that if we relax the conditions of Theorems 3.91, 3.95, 3.99 or 3.101, then statistical preference do es not hold in general. In particular, we replace the hyp othesis X $_{FSD}$ X_i by X $_{SP}$ X_i for some i, and we prove that X isnot necessarilythepreferred variable.

Example 3.104 *Considertheabsolutely continuousrandomvariables $X, X_1, \ldots, X_n$, whose density functions are givenby:*

$$\begin{aligned} f_X(t) &= I_{(2,3)} \\ f_{X_1}(t) &= 0.6\, I_{(1,2)}(t) + 0.4\, I_{(3,4)}(t). \\ f_{X_2}(t) &= I_{(2,3)} \\ f_{X_i}(t) &= I_{(0,1)} \text{ for any } i = 3, \ldots, n. \end{aligned}$$

It holds that X $_{SP}$ X_i for every $i = 1, \ldots, n$ and X $_{FSD}$ X_i for every $i = 2, \ldots, n$, but X $_{FSD}$ X_1. Moreover,

$$\begin{aligned} Q_n(X_1, [X, X_{-1}]) &= P(X_1 \quad (3,4)) = 0.4. \\ Q_n(X, [X_1, \ldots, X_n]) &= Q(X_2, [X, X_{-2}]). \\ Q_n(X_i, [X, X_{-i}]) &= 0 \text{ for any } i = 3, \ldots, n. \end{aligned}$$

Since thesum of these values is 1:

$$Q_n(X, [X_1, \ldots, X_n]) = Q(X_2, [X, X_{-2}]) = \frac{1}{2}(1 - Q_n(X_1, [X, X_{-i}])) = 0.3,$$

and therefore X_1 is not thepreferred randomvariable with respect to the general statistical preference.

Thus, Theorems 3.91, 3.95, 3.99 and 3.101 cannot be extended to any general situations.

3.3.4 General statistical preference vs m^{th} degree stochastic dominance

In the previous section we established conditions for first degree stochastic dominance to imply general statistical preference. Next we shall investigate the possible relationships between the m^{th} degree stochastic dominance and the general statistical preference.

Consider random variables $X, X_1, \dots, X_n$ and assume that $X \geq_{mSD} X_i$ ($m \geq 2$) for every $i = 1, \dots, n$. We shall study if under those conditions $X \ {}_{SP} [X_1, \dots, X_n]$. To see that this is not necessarily the case, consider the absolutely continuous random variables whose density functions are given by:

$$f_X(t) = I_{(5,6)}(t).$$
$$f_{X_1}(t) = 0.4\, I_{(0,1)}(t) + 0.6\, I_{(6,7)}(t).$$
$$f_{X_i}(t) = I_{(-1,0)}(t) \text{ for every } i = 2, \dots, n.$$

Then $X \geq_{mSD} X_i$ for every $i = 1, \dots, n$. In fact, $X \ {}_{FSD} X_i$ for every $i = 2, \dots, n$. However, X is not statistically preferred to $[X_1, \dots, X_n]$:

$$Q_n(X, [X_1, \dots, X_n]) = P(X > \max(X_1, \dots, X_n)) = P(X_1 \ (0, 1)) = 0.4.$$
$$Q_n(X_1, [X, X_j : j = 1]) = P(X_1 > \max(X, X_j : j = 1)) = P(X_1 \ (6, 7)) = 0.6.$$
$$Q_n(X_i, [X, X_j : j = i]) = 0 \quad \text{for any } i = 2, \dots, n.$$

Note that due to the definition of the density functions, the values of the relation Q_n are independent of the possible dependence among the random variables. Thus, we conclude that, for $m \geq 2$:

$$X \geq_{mSD} X_i \text{ for every } i = 1, \dots, n \text{ does not imply } X \ {}_{SP} [X_1, \dots, X_n].$$

Assume on the other hand that $X \ {}_{SP} [X_1, \dots, X_n]$ and let us investigate whether if $X \geq_{mSD} X_i$ for some $m \geq 1$. To see that this is not the case, consider the absolutely continuous random variables with density functions

$$f_X(t) = 0.4\, I_{(0,1)}(t) + 0.6\, I_{(2,3)}(t).$$
$$f_{X_i}(t) = I_{(1,2)}(t) \text{ for every } i = 1, \dots, n.$$

$X \ {}_{SP} [X_1, \dots, X_n]$, because:

$$Q_n(X, [X_1, \dots, X_n]) = P(X > \max(X_1, \dots, X_n)) = P(X \ (2, 3)) = 0.6.$$

However, X does not stochastically dominate X_i by the m^{th} degree for any $m \geq 1$, since $F_X(t) > F_{X_i}(t)$ for every $t \ (0, 1)$ and consequently $G^m_X(t) > G^m_{X_i}(t)$ for every $m \geq 2$ and $t \ (0, 1)$.

We conclude that $X \ _{\mathrm{SP}} [X_1, \ldots, X_n]$ do es not imply that exists $m \geq 1$ such that $X \geq_{\mathrm{mSD}} X_i$ for every $i = 1, \ldots, n$. This generalises Remark3.63, where we sawthat there is not a general relationship between the n^{th} degree sto chastic dominance and the pairwise statistical preference.

Remark 3.105 *Let us note that if $X, X_1, \ldots, X_n$ are $n+1$ random variablessuch that $X \ _{\mathrm{SP}} \max(X_1, \ldots, X_n)$ (respectively, $X \geq_{\mathrm{mSD}} \max(X_1, \ldots, X_n)$), then $X \ _{\mathrm{SP}} X_i$ (respectively, $X \geq_{\mathrm{mSD}} X_i$) for every $i = 1, \ldots, n$.*

To conclude this section, we presentthis result:

Prop osition 3.106 *Given $n+1$ real-valuedrandom variables $X, X_1, \ldots, X_n$, $X \ _{\mathrm{SP}} \max(X_1, \ldots, X_n)$ implies that $X \ _{\mathrm{SP}} [X_1, \ldots, X_n]$.*

Pro of Since $X \ _{\mathrm{SP}} \max(X_1, \ldots, X_n)$, it holdsthat

$$Q(X, \max(X_1, \ldots, X_n)) \geq Q(\max(X_1, \ldots, X_n), X).$$

Inparticular, by Lemma 2.20, we know that

$$P(X > \max(X_1, \ldots, X_n)) \geq P(\max(X_1, \ldots, X_n) > X),$$

since:

$$\begin{aligned}
P(X > \max(X_1, \ldots, X_n)) &\geq P(\max(X_1, \ldots, X_n) > X) \\
&= \underset{\substack{k=1,\ldots,n \\ 1 \leq i_1 < \ldots < i_k \leq n \\ i = i_1, \ldots, i_k}}{} P(X_i = X_{i_1} = \ldots = X_{i_k} > X, \max_{j=i_1,\ldots,i_k}(X_j)) \\
&\geq \underset{\substack{k=1,\ldots,n \\ 1 \leq i_1 < \ldots < i_k \leq n \\ i = i_1, \ldots, i_k}}{} \frac{1}{k+1} P(X_i = X_{i_1} = \ldots = X_{i_k} > X, \max_{j=i_1,\ldots,i_k}(X_j)).
\end{aligned}$$

Then:

$$\begin{aligned}
Q_n(X, [X_1, \ldots, X_n]) = &\underset{\substack{k=0,\ldots,n \\ 1 \leq i_1 < \ldots < i_k \leq n}}{} \frac{1}{k+1} P(X = X_{i_1} = \ldots = X_{i_k} > \max_{j=i_1,\ldots,i_k}(X_j)) \geq \\
&\underset{\substack{k=1,\ldots,n \\ 1 \leq i_1 < \ldots < i_k \leq n \\ i = i_1, \ldots, i_k}}{} \frac{1}{k+1} P(X_i = X_{i_1} = \ldots = X_{i_k} > X, \max_{j=i_1,\ldots,i_k}(X_j)) + \\
&\underset{\substack{k=1,\ldots,n \\ 1 \leq i_1 < \ldots < i_k \leq n \\ i = i_1, \ldots, i_k}}{} \frac{1}{k+1} P(X_i = X_{i_1} = \ldots = X_{i_k} > X, \max_{j=i_1,\ldots,i_k}(X_j)) \\
= &Q_n(X_i, [X, X_{-i}]).
\end{aligned}$$

Figure3.5 summarises some of the results of this section. Missin g arrows mean that an implication do es not hold in general, arrows with reference s m eans that such implication holdsin the conditions of such references, andarrowwithout reference means that such implication always holds.

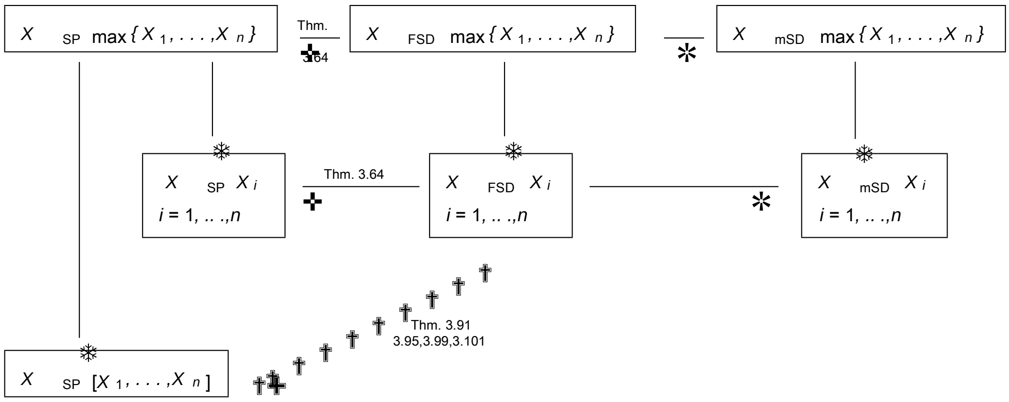


Figure 3.5: Relationships among firstand n^{th} degree sto chastic dominance, statistical preferenceand the generalstatistical preference.

3.4 Applications

In this section we present two possible applications of sto chastic orders. Onthe one hand, we apply sto chastic dominance and statistical preference for the comparison of fitness values, and on the other hand, we use the general statistical preference in decision makingproblems with linguistic variables.

3.4.1 Comparison of fitness values

Genetic algorithms are a p owerful to ol to p erform tasks such as generation of fuzzy rule bases, optimization of fuzzy rule bases, generation of memb ership functions, and tuning of memb ership functions (see [41]). All these tasks can be considered as optimization or search pro cesses. A geneticalgorithmgeneratesoradaptsa fuzzysystem, whichiscalled Genetic Fuzzy Systems (GFS, forshort) [42]. The use of GFS has been widely accepted, since these algorithms are robust and can search efficiently large solution spaces (see [213]).

Although in this context the linguistic granules or information are represented by fuzzy sets, the input data and the output results are usually crisp [87]. However, some recent papers (see [180, 181, 182, 183]) have dealt with fuzzy-valued data to learn and evaluate GFS. In that approach the function that quantifies the optimality of a solution in the genetic algorithm, that is, the fitness function, is fuzzy-valued. In particular, in [183], it has been considered that the fitness values are unknown, and that interval valued information is available. The computed fitness value is used by the genetic algorithm module to produce the next population of individuals. In this context some kind of order between two fitness values is necessary if we want to determine whether one individual precedes the other. Since the information about the fitness values is imprecise and is given by means of intervals, a procedure for comparing two intervals is required. Initially, these procedures were based on estimating and comparing two probabilities [183]. In this section we consider statistical preference as a more flexible tool for the comparison of intervals.

Thus, in this section we study of these concepts in connection with the comparison of two intervals, that represent imprecise information about the fitness values of two Knowledge Bases. In particular, we shall make no assumptions about the joint distribution of the two fitness values and shall use then the uniform distribution. This is not an artificial requirement, and it has been considered in many situation as a consequence of lack of information (see, for instance, [183, 197]). When this distribution is considered, we obtain the specific expression of the associated probabilistic and fuzzy relations. We also consider the situation where we have some additional information about the distribution of the fitness, that we model that by means of beta distributions. For these two cases, we consider three possible situations between the intervals: independence, comonotonicity and countermonotonicity.

Usual comparison methods

Let us consider two fitness values θ_1 and θ_2 of two KBs, that is, the mean squared errors of these two KBs on the training set. In many situations, θ_1 and θ_2 are unknown, but we have some imprecise information about them, that we model by means of two intervals that include them. These intervals can be obtained by means of a fuzzy generalisation of the mean squared errors (for a more detailed explanation, see Sections 4 and 5 in [183]) and they will be denoted by FMSE_1 and FMSE_2, respectively. The comparison of this two intervals is needed in order to choose the predecessor and the successor.

Let us introduce the usual methods that can be found in the literature for the comparison of such intervals. We shall propose statistical preference as an alternative method and investigate the relationships between all the possibilities.

Let us start with the *strong dominance* that was considered in [116].Inthatcase, if these two intervals are disjoint, then wehave notany problem todetermine the preferred interval andtherefore the decision is trivial. Theproblemarises whentheintersectionis non-empty,since theintervals are incomparable.

Definition 3.107 *Considerthe fitness θ_1 and θ_2 withassociated intervalsFMSE $_1$ = $[a_1,b_1]$ and $FMSE_2 = [a_2,b_2]$,respectively. It holdsthat:*

- *If $b_2 < a_1$, then θ_1 ispreferred to θ_2 with respect to the strong dominance, denoted by $\theta_1 \ _{sd}\ \theta_2$.*
- *If $b_1 < a_2$, then θ_2 ispreferred to θ_1 with respect to the strong dominance, denoted by $\theta_2 \ _{sd}\ \theta_1$.*
- *Otherwise, θ_1 and θ_2 are incomparable.*

This metho d is to o restrictive, since it can be used only in very particular cases. An attempt to solve this problem is to use the first degree sto chastic dominance, that introduces prior knowledge ab out the probability distribution of the fitness.

In particular, if we assume thatthe fitness follows a uniform distribution (as in [197]), then:

$$\theta_1 \ _{FSD}\ \theta_2 \quad a_1 \geq a_2 \text{ and } b_1 \geq b_2,$$

with at least one of the inequalities strict. In parti cular, if θ_1 strong dominates θ_2, then $\theta_1 \ _{FSD}\ \theta_2$ regardlesson thedistributionofthe fitness.

Nevertheless, first degree sto chastic dominance, as we have already noticed during this memory, do es not solve all the problems of strong domi nance,since, for instance, incomparability is also allowed.

Another metho d, called *method of the probabilistic prior*, was prop osed in [183]. As first degree sto chastic dominance, it is based on a prior knowledge ab out the probability distribution of the fitness, $P(\theta_1,\theta_2)$.

Definition 3.108 *Considerthe fitness θ_1 and θ_2 withassociated intervalsFMSE $_1$ = $[a_1,b_1]$ and $FMSE_2 = [a_2,b_2]$. Then, θ_1 is considered tobe preferred to θ_2 withrespect to the probabilistic prior, and is denoted by $\theta_1 \ _{pp}\ \theta_2$, if and only if*

$$\frac{P(\theta_1 > \theta_1)}{P(\theta_1 \leq \theta_2)} > 1. \tag{3.30}$$

If $P\{\theta_1 \leq \theta_2\} = 0$, the ration inEquation (3.30) *is notdefined, bu t it is assumed that $\theta_1 \ _{pp}\ \theta_2$.*

Remark 3.109 *Recall that from Equation* (3.30) *we derive that* $\theta_1 \ {}_{pp}\ \theta_2$ *if and only if:*

$$P(\theta_1 > \theta_1) > P(\theta_1 \leq \theta_2).$$

Thus, the probability prior is equivalent to the probability dominance, with the strict version, considered in Remark 2.22, with $\beta = 0.5$.

Even though these methods allow to compare a wider class of random intervals than the strong dominance, as we said in Remark 2.22 they have an important drawback: they allow for incomparability. In particular, whenever $P(\theta_1 = \theta_2) \geq 0.5$, θ_1 and θ_2 would be incomparable.

Then, it seems natural to consider statistical preference as a method for the comparison of fitness for two main reasons: avoid incomparability and graduate the preference. Also, as we already commented in Subsection 2.1.2, the probabilistic relation Q can be transformed into a fuzzy relation.

Let us study some relationships among strong dominance, first degree stochastic dominance, probabilistic prior and statistical preference.

Proposition 3.110 *Given two fitness* θ_1 *and* θ_2 *with associated intervals* $FMSE_1 = [a_1, b_1]$ *and* $FMSE_2 = [a_2, b_2]$*, it holds that:*

- $\theta_1 \ {}_{sd}\ \theta_2$ *implies* $\theta_1 \ {}_{FSD}\ \theta_2$.
- $\theta_1 \ {}_{sd}\ \theta_2$ *implies* $\theta_1 \ {}_{pp}\ \theta_2$.
- $\theta_1 \ {}_{pp}\ \theta_2$ *implies* $\theta_1 \ {}_{SP}\ \theta_2$.
- *If* θ_1 *and* θ_2 *are independent,* $\theta_1 \ {}_{FSD}\ \theta_2$ *implies* $\theta_1 \ {}_{pp}\ \theta_2$.

Proof

- The proof of the first item is based on the fact that $\theta_1 \ {}_{sd}\ \theta_2$ implies

  $$\min FMSE_1 = a_1 > b_2 \max FMSE_2,$$

  and consequently $\theta_1 \ {}_{FSD}\ \theta_2$ regardless on the distributions of $FMSE_i$, $i = 1, 2$.
- If $\theta_1 \ {}_{sd}\ \theta_2$, then $\{(\theta_1, \theta_2) : \theta_1 \leq \theta_2\} =$, and consequently $\theta_1 \ {}_{pp}\ \theta_2$.
- If $\theta_1 \ {}_{pp}\ \theta_2$, then $P(\theta_1 > \theta_2) > P(\theta_1 \leq \theta_2)$, that implies $Q(X, Y) > Q(Y, X)$. However, since Q is a probabilistic relation, this means that $Q(X, Y) > \frac{1}{2}$, and thus $\theta_1 \ {}_{SP}\ \theta_2$.

- If the intervals are indep endent, then $P(\theta_1 = \theta_2) = 0$, and consequently $\theta_1 \succeq_{pp} \theta_2$ ifand only if

$$P(\theta_1 > \theta_2) > P(\theta_1 < \theta_2).$$

Thus, b oth the probabilis tic prior and statis tical preference are equivalent in this context. Thus, if $\theta_1 \succeq_{FSD} \theta_2$, applying Theorem 3.64, $\theta_1 \succeq_{SP} \theta_2$, and conse quently the preference with resp ect to the probabilistic prior metho d also hold. ■

Thus there is a relationship b etween the probabilistic prior and the sto chastic order when the intervals are indep endent.However, such relationship do es not hold for comonotonic and countermonotonic intervals , as we show next:

Example 3.111 *Consider θ_1 distributed inthe interval $[1, 2]$ and θ_2 distributed inthe interval $[0, 2]$ We consider that $FMSE_1$ fol lows an uniform distribution and the distribution of $FMSE_2$ is definedby the densityfunction:*

$$f(x) = \begin{cases} \frac{1}{11} & \text{if } 0 < x < 1.1, \\ 1 & \text{if } 1.1 < x < 2, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $\theta_1 \succeq_{FSD} \theta_2$. Assume that bothintervals arecomonotonic. UsingEquation (3.6) we can compute $P(\theta_1 = \theta_2)$:

$$P(\theta_1 = \theta_2) = \int_{[1.1,2]} f_X(x)dx = 0.9.$$

Thus, both intervals areincomparable withrespect tothe probabilistic prior.

Assume now that they are countermonotonic. Using Equation (3.7)we obtainthat

$$Q(\theta_1, \theta_2) = F_Y(1.5) = 0.5.$$

Thus, $\theta_1 \succeq_{SP} \theta_2$, and consequently, using Proposition 3.110, $\theta_1 \succeq_{pp} \theta_2$.

Table 3.6summarises the generalrelationships wehave seen during this section.

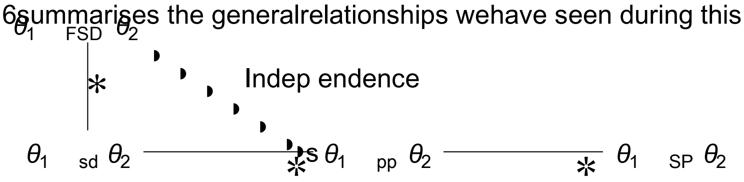


Figure 3.6: Summary of the relationships b etween strong dominance, first degree sto chastic dominance, probabilistic prior and statistical preference given in Prop osition 3.110.

Expression of the probabilistic relation for the comparison of fitness values

In this section we will apply statistical preference to the comparison of fitness values.

Uniform case Let us consider again an uniform distribution, that is, no prior information about the distribution over the observed interval, as in [197], and let us search for an expression of the probabilistic relation Q so as to characterise the statistical preference.

Thus, $FMSE_1 = [a_1, b_1]$ and $FMSE_2 = [a_2, b_2]$ will denote now two intervals where we know the fitness θ_1 and θ_2 of two KBs are included. Let us assume a uniform distribution on each of them. We will consider again three possible ways to obtain the joint distribution: an assumption of independence, that is, being coupled by the product, and the extreme cases where they are coupled by the minimum or the Łukasiewicz copulas. In these three cases we will obtain the condition on the parameters to assure the statistical preference of the interval $FMSE_1$ to the interval $FMSE_2$. To do that, the expression of the probabilistic relation will be an essential part of the proof.

First of all, recall the result the comparison of independent uniform distributions was already studied in Proposition 3.71: if $FMSE_1 = [a_1, b_1]$ and $FMSE_2 = [a_2, b_2]$ be two uniformly distributed intervals which represent the information we have about the fitness θ_1 and θ_2 of two KBs, and the joint distribution is obtained by means of the product copula, then the probabilistic relation $Q(\theta_1, \theta_2)$ takes the following value:

$$Q(\theta_1,\theta_2) = \begin{cases} 1 - \frac{(b_1-a_2)^2}{2(b_1-a_1)(b_2-a_2)} & \text{if } a_1 \le a_2 < b_1 \le b_2. \\ 1 - \frac{(b_2-a_1)^2}{2(b_1-a_1)(b_2-a_2)} & \text{if } a_2 \le a_1 < b_2 \le b_1. \\ \frac{2b_1-a_2-b_2}{2(b_1-a_1)} & \text{if } a_1 \le a_2 < b_2 \le b_1. \\ \frac{b_1+a_1-2a_2}{2(b_2-a_2)} & \text{if } a_2 \le a_1 < b_1 \le b_2. \end{cases}$$

These are the conditions under which $\theta_1 \succ_{SP} \theta_2$:

$$\begin{cases} \text{Always} & \text{if } a_1 \le a_2 < b_1 \le b_2. \\ \text{Never} & \text{if } a_2 \le a_1 < b_2 \le b_1. \\ a_1 + b_1 \ge b_2 + a_2 & \text{if } a_1 \le a_2 < b_2 \le b_1 \\ & \text{or } a_2 \le a_1 < b_1 \le b_2. \end{cases}$$

Let us now study the comonotonic case.

Proposition 3.112 *Let $FMSE_1 = [a_1, b_1]$ and $FMSE_2 = [a_2, b_2]$ be two uniformly distributed intervals representing the available information on the different fitness θ_1 and θ_2 of two KBs. If the joint distribution is obtained by means of the minimum copula, the* *probabilistic relation* $Q(\theta_1,\theta_2)$ *takes the fol lowing value:*

$$Q(\theta_1,\theta_2)=\begin{cases} 0 & \text{if } a_1 \leq a_2 < b_1 \leq b_2. \\ \frac{b_1-b_2}{b_1+a_2-a_1-b_2} & \text{if } a_1 \leq a_2 < b_2 < b_1. \\ \frac{a_1-a_2}{b_2-a_2-b_1+a_1} & \text{if } a_2 < a_1 < b_1 \leq b_2. \\ 1 & \text{if } a_2 < a_1 < b_2 \leq b_1. \end{cases}$$

Thus, $\theta_1 \ _{SP}\ \theta_2$ *if andonly if:*

$$\begin{cases} \textit{Never} & \text{if } a_1 \leq a_2 < b_1 \leq b_2. \\ \textit{Always} & \text{if } a_2 < a_1 < b_2 \leq b_1. \\ a_1+b_1 > a_2+b_2 & \textit{otherwise.} \end{cases}$$

Then, the condition is equivalent to have a greater expectation.

Pro of The expression of the probabilistic relation can b e ob tained using Equation (3.6), and taking into account that $P(\theta_1=\theta_2)=0$, since the asso ciated cumulative distribution coincide at most in one point.

First and second scenarios of the are trivial. In thethird scenario, if $a_1 \leq a_2 < b_2 \leq b_1$ itholds that:

$$\theta_1 \ _{SP}\ \theta_2 \quad \frac{b_1-b_2}{b_1+a_2-a_1-b_2} > \frac{1}{2} \quad a_1+b_1 > a_2+b_2.$$

The condition for $a_2 \leq a_1 < b_1 \leq b_2$ can b e sim ilarly obtained. ∎

Finally, letus studythecountermonotonic case.

Prop osition 3.113 *Let* $FMSE_1=[a_1,b_1]$ *and* $FMSE_2=[a_2,b_2]$ *be twouniformly distributed intervals whichrepresentthe information we have about the fitness* θ_1 *and* θ_2 *of two KBs. Ifthejoint distributionis obtainedbymeans oftheŁukasiewiczcopula, then the probabilistic relation is given by:*

$$Q(\theta_1,\theta_2)=\frac{b_1-a_2}{b_2-a_2+b_1-a_1}.$$

In addition, $\theta_1 \ _{SP}\ \theta_2$ *if andonly if:*

$$\begin{cases} \textit{Never} & \text{if } a_1 \leq a_2 < b_1 \leq b_2. \\ a_1+b_1 \geq a_2+b_2 & \text{if } a_1 \leq a_2 < b_2 < b_1. \\ a_1+b_1 \geq a_2+b_2 & \text{if } a_2 < a_1 < b_1 \leq b_2. \\ \textit{Always} & \text{if } a_2 < a_1 < b_2 \leq b_1. \end{cases}$$

Proof The expression of the probabilistic relation can be obtained using Equation (3.7), and taking into account that the point u such that $F_{\theta_1}(u) + F_{\theta_2}(u) = 1$ equals: $u = \frac{b_2 b_1 - a_1 a_2}{b_2 - a_2 + b_1 - a_1}$.

The first and fourth scenarios of the second part are easy, since there they are ordered by means of the stochastic order. In the first scenario it holds that $F_{\theta_1}(u) > F_{\theta_2}(u)$, and consequently

$$Q(\theta_1, \theta_2) < Q(\theta_2, \theta_1),$$

and then $\theta_1 \ _{SP}\ \theta_2$. Similarly, we obtain that in the fourth scenario $\theta_1 \ _{SP}\ \theta_2$.

For the second and third scenarios, it is enough to compare de expression of the probabilistic relation with $\frac{1}{2}$. ∎

Beta case We now assume that more information about the fitness values may be available. If it is know that some values of the interval are more feasible than others, the uniform distribution is not a good model any more. If we assume that the closer we are to one extreme of the interval the more feasible the values are, beta distributions become more appropriate to model the fitness values. As we made in Subsection 3.2.6, we focus on this situation: beta distributions such that one of the parameters is 1.

As we already said, the density of a beta distribution $\beta(p, q)$ is given by Equation (3.17). However, it is possible to define a beta distribution on every interval $[a, b]$ (it is denoted by $\beta(p, q, a, b)$) The associated density function is:

$$f(x) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \frac{(x-a)^{p-1}(b-a)^{q-1}}{(b-a)^{p+q-1}},$$

for any x $[a, b]$ and zero otherwise. Next, we will focus on two particular cases. In the first one we will assume that the closer the value is to a_i, the more feasible the value is. In the second case, we will assume the opposite: that the closer the value is to b_i, the more feasible the value is. In terms of density functions, these two cases correspond to strictly decreasing and strictly increasing density functions. We will consider the intervals FMSE$_i$ follows a distribution $\beta(p, 1, a_i, b_i)$, for $i = 1, 2$, where p will be an integer greater than 1. Independently of where the weight of the distribution is, we shall consider three possibilities concerning the relationship between the fitness values: independence, comonotonicity and countermonotonicity. If intervals satisfy one of the following conditions:

$$a_1 \leq a_2 < b_1 \leq b_2 \text{ or } a_2 \leq a_1 < b_2 \leq b_1,$$

we have seen in the previous section that, since they are ordered with respect to the stochastic order, the study of the statistical preference becomes trivial. For this reason we will assume the intervals to satisfy the condition $a_1 \leq a_2 < b_2 \leq b_1$ (the case $a_2 \leq a_1 < b_1 \leq b_2$ can be solved by symmetry).

Proposition 3.114 *Let us consider the different fitness values θ_1 and θ_2 with associated intervals $FMSE_i \equiv [a_i, b_i]$ following a distribution $\beta(p, 1, a_i, b_i)$, where $a_1 \leq a_2 < b_2 \leq b_1$. Then:*

$$Q^P(\theta_1,\theta_2) = p \sum_{k=0}^{p-1} \binom{p-1}{k} \frac{(a_2 - a_1)^{p-k-1}(b_2 - a_2)^{k-1}}{(b_1 - a_1)^p(p + k + 1)} + \left(\frac{b_2 - a_1}{b_1 - a_1}\right)^p - \left(\frac{a_2 - a_1}{b_1 - a_1}\right)^p,$$

$$Q^M(\theta_1,\theta_2) = 1 - \left(\frac{t - a_1}{b_1 - a_1}\right)^p,$$

$$Q^L(\theta_1,\theta_2) = \left(\frac{z - a_2}{b_2 - a_2}\right)^p,$$

where $t = \frac{a_1 b_2 - a_2 b_1}{b_2 - a_2 - b_1 + a_1}$ *and z is the point in $[a_2, b_2]$ such that*

$$\left(\frac{z - a_1}{b_1 - a_1}\right)^p + \left(\frac{z - a_2}{b_2 - a_2}\right)^p = 1,$$

and Q^P, Q^M and Q^L denotes the probabilistic relation when the random variables are coupled by the product, the minimum and the Łukasiewicz operators, respectively.

Proof Let us begin by computing the expression of $Q^P(\theta_1,\theta_2)$. Since they are independent and continuous, $P(\theta_1 = \theta_2) = 0$. Then:

$$Q^P(\theta_1,\theta_2) = P(\theta_1 > \theta_2) = P(\theta_1 \in [b_2, b_1]) + P(b_2 > \theta_1 > \theta_2).$$

Let us compute each one of the previous probabilities:

$$P(\theta_1 \in [b_2, b_1]) = \int_{a_2}^{b_2} p\frac{(x - a_1)^{p-1}}{(b_1 - a_1)^p}\,dx = \left(\frac{b_2 - a_1}{b_1 - a_1}\right)^p - \left(\frac{a_2 - a_1}{b_1 - a_1}\right)^p.$$

$$\begin{aligned}
P(b_2 > \theta_1 > \theta_2) &= \int_{a_2}^{b_2}\int_{a_2}^{x} p^2 \frac{(x - a_1)^{p-1}}{(b_1 - a_1)^p}\frac{(y - a_2)^{p-1}}{(b_2 - a_2)^p}\,dy\,dx \\
&= \int_{a_2}^{b_2} p\frac{(x - a_1)^{p-1}}{(b_1 - a_1)^p}\left(\frac{x - a_2}{b_2 - a_2}\right)^p dx.
\end{aligned}$$

Taking $z = \frac{x - a_2}{b_2 - a_2}$, the previous expression becomes:

$$\begin{aligned}
P(b_2 > \theta_1 > \theta_2) &= p\int_0^1 \frac{\left((b_2 - a_2)z + a_2 - a_1\right)^{p-1}}{(b_1 - a_1)^p} z^p \frac{dz}{b_2 - a_2} \\
&= p\int_0^1 \frac{z^p}{(b_2 - a_2)(b_1 - a_1)^p}\sum_{k=0}^{p-1}\binom{p-1}{k}((b_2 - a_2)z)^k(a_2 - a_1)^{p-1-k}\,dz \\
&= \frac{p}{(b_1 - a_1)^p}\sum_{k=0}^{p-1}\binom{p-1}{k}\frac{(a_2 - a_1)^{p-k-1}(b_2 - a_2)^{k-1}}{p + k + 1}.
\end{aligned}$$

Making the sum of the two probabilities, we obtain the value of $Q(\theta_1,\theta_2)$.

Next, assume that θ_1 and θ_2 are comonotonic. Since $\{x: F_{\theta_1}(x) = F_{\theta_2}(x)\} = \emptyset$, applying Equation (3.6) we deduce that

$$Q^M(\theta_1,\theta_2) = \int_{x:F_{\theta_1}(x)<F_{\theta_2}(x)} p\frac{(x-a_1)^{p-1}}{(b_1-a_1)^p}\,dx.$$

Moreover, $\{x: F_{\theta_1}(x) < F_{\theta_2}(x)\} = (t, b_1]$, where t is the point satisfying:

$$F_{\theta_1}(t) = F_{\theta_2}(t) \Leftrightarrow \left(\frac{t-a_1}{b_1-a_1}\right)^p = \left(\frac{t-a_2}{b_2-a_2}\right)^p$$
$$\Leftrightarrow t(b_2-a_2) - a_1(b_2-a_2) = t(b_1-a_1) - a_2(b_1-a_1)$$
$$\Leftrightarrow t = \frac{a_1b_2-a_2b_1}{b_2-a_2-b_1+a_1}.$$

Then:

$$Q^M(\theta_1,\theta_2) = \int_t^{b_1} p\frac{(x-a_1)^{p-1}}{(b_1-a_1)^p}\,dx = 1 - \left(\frac{t-a_1}{b_1-a_1}\right)^p.$$

Finally, assume that θ_1 and θ_2 are countermonotonic. By Equation (3.7),

$$Q^L(\theta_1,\theta_2) = F_{\theta_2}(z) = \left(\frac{z-a_2}{b_1-a_1}\right)^p,$$

where z satisfies that:

$$F_{\theta_1}(z) + F_{\theta_2}(z) = 1 \Leftrightarrow \left(\frac{z-a_1}{b_1-a_1}\right)^p + \left(\frac{z-a_2}{b_2-a_2}\right)^p = 1. \qquad \blacksquare$$

Proposition 3.115 *Let us consider the different fitness values θ_1 and θ_2 with associated intervals $FMSE_1 = [a_1, b_1]$ and $FMSE_2 = [a_2, b_2]$ following the distribution $\beta(1, q, a_i, b_i)$, where $a_1 \leq a_2 < b_2 \leq b_1$. Then*

$$Q^P(\theta_1,\theta_2) = q\sum_{k=0}^{q-1}\binom{q-1}{k}\frac{(b_1-b_2)^k(b_2-a_2)^{q-k-2}}{(b_1-a_1)^q(q+k+1)} + \left(\frac{b_1-a_2}{b_1-a_1}\right)^q,$$
$$Q^M(\theta_1,\theta_2) = 1 - \left(\frac{b_1-t}{b_1-a_1}\right)^q,$$
$$Q^L(\theta_1,\theta_2) = 1 - \left(\frac{b_1-z}{b_1-a_1}\right)^p,$$

where $t = \frac{a_1b_2-b_1a_2}{b_2-a_2-b_1+a_1}$ and z is the point in $[a_2, b_2]$ such that

$$\frac{(b_1-x)^q}{(b_1-a_1)^{q-1}} + \frac{(b_2-x)^q}{(b_2-a_2)^{q-1}} = 1,$$

and Q^P, Q^M and Q^L denotes the probabilistic relation when the random variables are coupled by the product, the minimum and the Łukasiewicz operators, respectively.

Pro of We begin by computing the expression of $Q^P(\theta_1,\theta_2)$. Again, sincethey are indep endent and continuous $P(\theta_1=\theta_2)=0$,and then:

$$Q^P(\theta_1,\theta_2) = P(\theta_1>\theta_2) = P(\theta_1 \in [b_2,b_1]) + P(b_2>\theta_1>\theta_2).$$

Letus compute each onethe the previous probabilities:

$$P(\theta_1 \in [b_2,b_1]) = \int_{b_2}^{b_1} q\frac{(b_1-x)^{q-1}}{(b_1-a_1)^q}dx = \left(\frac{b_1-b_2}{b_1-a_1}\right)^q.$$

$$\begin{aligned}
P(b_2>\theta_1>\theta_2) &= \int_{a_2}^{b_2}\int_{a_2}^{x} q^2\frac{(b_1-x)^{q-1}}{(b_1-a_1)^q}\frac{(b_2-y)^{q-1}}{(b_2-a_2)^q}dy\,dx \\
&= \int_{a_2}^{b_2} q\frac{(b_1-x)^{q-1}}{(b_1-a_1)^q}\left(1-\left(\frac{b_2-x}{b_2-a_2}\right)^q\right)dx \\
&= \int_{a_2}^{b_2} q\frac{(b_1-x)^{q-1}}{(b_1-a_1)^q}dx - \int_{a_2}^{b_2} q\frac{(b_1-x)^{q-1}}{(b_1-a_1)^q}\left(\frac{b_2-x}{b_2-a_2}\right)^q dx \\
&= \frac{b_1-a_2}{b_1-a_1} - \frac{b_1-b_2}{b_1-a_1} - \int_{a_2}^{b_2} q\frac{(b_1-x)^{q-1}}{(b_1-a_1)^q}\left(\frac{b_2-x}{b_2-a_2}\right)^q dx.
\end{aligned}$$

Taking $z=\frac{b_2-x}{b_2-a_2}$, the last integral becomes:

$$\begin{aligned}
P(b_2>\theta_1>\theta_2) &= \int_0^1 qz^q\frac{(b_1-b_2+z(b_2-a_2))^{q-1}}{(b_1-a_1)^q}\frac{dz}{b_2-a_2} \\
&= q\int_0^1 \frac{z^q}{(b_2-a_2)(b_1-a_1)^q}\sum_{k=0}^{q-1}\binom{q-1}{k}((b_1-b_2)z)^k(b_2-a_2)^{q-k-1}dz \\
&= q\sum_{k=0}^{q-1}\binom{q-1}{k}\frac{(b_1-b_2)^k(b_2-a_2)^{q-k-2}}{(b_1-a_1)^q(q+k+1)}\frac{1}{q+k+1}.
\end{aligned}$$

Making the sum ofthethree terms,we obtain the expression of $Q^P(\theta_1,\theta_2)$.

Consider now the fitness to be comonotonic.Then, since $\{x : F_{\theta_1}(x) = F_{\theta_2}(x)\} = \emptyset$, the expression of the probabilistic relation given in Eq.(3.6) becomes:

$$Q^M(\theta_1,\theta_2) = \int_{x:F_{\theta_1}(x)<F_{\theta_2}(x)} q\frac{(b_1-x)^{q-1}}{(b_1-a_1)^q}dx.$$

Then, $\{x : F_{\theta_1}(x) < F_{\theta_2}(x)\} = (t, b_1]$,where:

$$\begin{aligned}
F_{\theta_1}(t) = F_{\theta_2}(t) &\Leftrightarrow 1-\left(\frac{b_1-t}{b_1-a_1}\right)^q = 1-\left(\frac{b_2-t}{b_2-a_2}\right)^q \\
&\Leftrightarrow \frac{b_1-t}{b_1-a_1} = \frac{b_2-t}{b_2-a_2} \Leftrightarrow t = \frac{a_1b_2-b_1a_2}{b_2-a_2-b_1+a_1}.
\end{aligned}$$

Then:

$$Q^M(\theta_1,\theta_2) = \int_t^{b_1} q\frac{(b_1-x)^{q-1}}{(b_1-a_1)^q}\mathrm{d}x = \left(\frac{b_1-x}{b_1-a_1}\right)^q.$$

Finally, assume that θ_1 and θ_2 are countermonotonic. Then, $Q^L(\theta_1,\theta_2) = F_{\theta_2}(z)$, where z satisfies:

$$F_{\theta_1}(z) + F_{\theta_2}(z) = 1 \quad 1 - \left(\frac{b_1-x}{b_1-a_1}\right)^q + 1 - \left(\frac{b_2-x}{b_2-a_2}\right)^q = 1$$
$$\left(\frac{b_1-x}{b_1-a_1}\right)^q + \left(\frac{b_2-x}{b_2-a_2}\right)^q = 1.$$

Remark 3.116 *In order to prove the previous result it is not possible to fol low the procedu re of Proposition 3.78.There, we used the fol lowing property:*

$$X \equiv \beta(p,1) \quad 1 - X \equiv \beta(1,p).$$

Then, since $Q(X,Y) = Q(1-Y, 1-X)$ *(see Proposit ion 3.3), the case of* $q=1$ *was solved using the case* $p=1$. *In thecase ofgeneral beta distributions, it holds that:*

$$X \equiv \beta(p,1,a,b) \quad (b-a) - X \equiv \beta(1,p,a,b).$$

The problem isthat $Q(X,Y) = Q((b_2-a_2)-Y, (b_1-a_1)-X)$, *andthereforethis kind of procedure is not possible.*

Remark 3.117 *Note that for beta distribution it is not possible to obtain a simpler characterization of the statistical preference like the one foruniform distributions.*

To conclude this section, let us present anexample where we show how the values of the probabilistic relation changewhen we vary thevalue of p.

Example 3.118 *Considerthe fitnessvalues* θ_1 *and* θ_2 *withassociatedvalues* $FMSE_1 = [a_1,b_1]$ *and* $FMSE_2 = [a_2,b_2]$, *where* $a_1 \leq a_2 < b_2 \leq b_1$, *and let assume they fol low the beta distribution* $\beta(p,1,a_i,b_i)$. *Consider* $a_1 = 1, b_1 = 4, a_2 = 2$ *and* $b_2 = 3$. *Table 3.5 shows thevalues of theprobabilistic relation when* p *moves from 1 to 5, whereit is possible to see that* θ_1 *and* θ_2 *areequivalent when* $p=1$, *but* θ_1 *ispreferred to* θ_2 *when* $p \geq 2$. *Moreover,the greater the value of* p, *the stronger the preference of* θ_1 *over* θ_2.

Consider now different values of the intervals: $a_1 = 0.7, b_1 = 1.4, a_2 = 0.8$ *and* $b_2 = 1.2$. *In this case, alt hough* $[a_2,b_2]$ $[a_1,b_1]$ *as inthe previousexample, the difference between* b_1 *and* b_2 *isgreater than* a_1 *and* a_2. *The results aresummarisedin Table 3.6. There, wecan seethat inthe threecases,* θ_1 $_{SP}$ θ_2 *for any* $p \geq 1$. *Furthermore, the greater the value ofp, thestrongerthe preferenceof* θ_1 *over* θ_2. *InFigure3.7 we cansee how thevalues of Qvary we changethe value of the parameter* p *from 1 to 10.*

p	Q^P	Q^M	Q^L
1	0.5	0.5	0.5
2	0.6853	0.75	0.64
3	0.7945	0.875	0.7436
4	0.8644	0.9375	0.8208
5	0.9101	0.9688	0.8766

Table 3.5: Degrees of preference for the different values of the parameter p for FMSE$_1$ = [1, 4] and FMSE$_2$ = [2, 3].

p	Q^P	Q^M	Q^L
1	0.5715	0.6667	0.5455
2	0.7076	0.8889	0.64
3	0.7936	0.9630	0.7192
4	0.8533	0.9877	0.7852
5	0.8955	0.9959	0.8384

Table 3.6: Degrees of preference for the different values of the parameter p for FMSE$_1$ = [0.7, 1.4] and FMSE$_2$ = [0.8, 1.2]

3.4.2 General statistical preference as a tool for linguistic decision making

As we have seen, general statistical preference was introduced as a method that allows for the comparison of more than two random variables. As an illustration of the utility of this method we can consider a decision making problem with linguistic utilities. We consider the example of product management given in [123, Section 8]: a company seeks to plan its production strategy for the next year, and they consider six possible alternatives:

- A_1 : Create a new product for very high-income customers.
- A_2 : Create a new product for high-income customers.
- A_3 : Create a new product for medium-income customers.
- A_4 : Create a new product for low-income customers.
- A_5 : Create a new product suitable for all customers.
- A_6 : Do not create a new product.

Due to the large uncertainty, the three experts of the company are not able to draw the information about the impact of each alternative in a numerical way, and for this reason

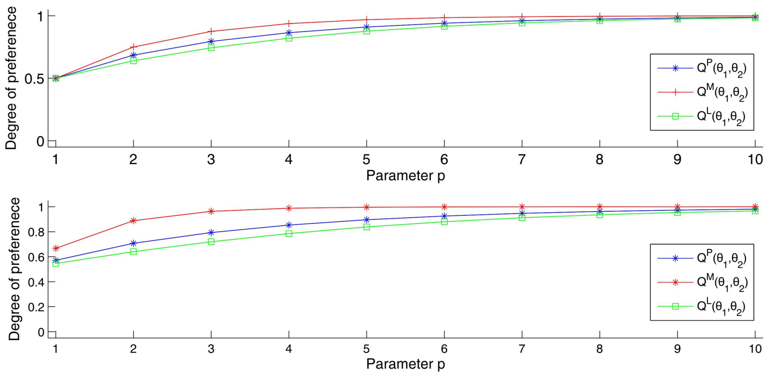


Figure 3.7: Values of the probabilistic relation for different values of p. The above picture corresponds to intervals $[a_1, b_1] = [1, 4]$ and $[a_2, b_2] = [2, 3]$, and the picture below corresponds to intervals $[a_1, b_1] = [0.7, 1.4]$ and $[a_2, b_2] = [0.8, 1.2]$

they express the utility based on a seven linguistic scale $S = \{s_1, \ldots, s_7\}$, where:

s_1 :	None	s_5 :	High
s_2 :	Very low	s_6 :	Very high
s_3 :	Low	s_7 :	Perfect
s_4 :	Medium		

Note that the three experts have not the same influence in the company, and its importance is given by the weight vector $(0.2, 0.4, 0.4)$. Moreover, since the decision of each expert depends on the economic situation of the following year, six scenarios are considered:

N_1 :	Very bad	N_4 :	Regular-Good
N_2 :	Bad	N_5 :	Good
N_3 :	Regular-Bad	N_6 :	Very good

The experts assume the following weighting vector for these scenarios:

$$W = (0.1, 0.1, 0.1, 0.2, 0.2, 0.3).$$

Finally, the preferences of each expert are given in Tables 3.7, 3.8 and 3.9.

Although in [123] this problem was solved by means of a particular type of aggregation operators, we propose to use the general statistical preference. For any expert e_i, $i = 1, 2, 3$, we can compute the preference degree of the alternative A_j over the others

	N_1	N_2	N_3	N_4	N_5	N_6
A_1	s_2	s_1	s_4	s_6	s_7	s_5
A_2	s_1	s_3	s_5	s_5	s_6	s_6
A_3	s_3	s_4	s_4	s_4	s_4	s_7
A_4	s_2	s_5	s_6	s_4	s_2	s_5
A_5	s_1	s_3	s_4	s_5	s_6	s_6
A_6	s_6	s_5	s_5	s_4	s_2	s_2

Table 3.7: Linguistic payoff matrix-Exp ert 1.

	N_1	N_2	N_3	N_4	N_5	N_6
A_1	s_3	s_1	s_3	s_5	s_6	s_6
A_2	s_1	s_3	s_4	s_5	s_6	s_6
A_3	s_3	s_4	s_5	s_4	s_3	s_7
A_4	s_3	s_4	s_5	s_4	s_2	s_4
A_5	s_2	s_3	s_4	s_6	s_6	s_6
A_6	s_7	s_6	s_4	s_3	s_2	s_2

Table 3.8: Linguistic payoff matrix-Exp ert 2.

	N_1	N_2	N_3	N_4	N_5	N_6
A_1	s_1	s_2	s_3	s_5	s_7	s_6
A_2	s_2	s_3	s_4	s_4	s_5	s_6
A_3	s_3	s_4	s_6	s_4	s_3	s_7
A_4	s_2	s_4	s_6	s_4	s_2	s_4
A_5	s_1	s_3	s_4	s_5	s_6	s_6
A_6	s_6	s_6	s_5	s_3	s_2	s_3

Table 3.9: Linguistic payoff matrix-Exp ert 3.

A_{-j}, and we obtain the following values:

$$Q(A_1, [A_{-1}] \mid e_1) = P(N_4) + P(N_5) = 0.4.$$
$$Q(A_2, [A_{-2}] \mid e_1) = 0.$$
$$Q(A_3, [A_{-3}] \mid e_1) = P(N_6) = 0.3.$$
$$Q(A_4, [A_{-4}] \mid e_1) = \tfrac{1}{2}P(N_2) + P(N_3) = 0.15.$$
$$Q(A_5, [A_{-5}] \mid e_1) = 0.$$
$$Q(A_6, [A_{-6}] \mid e_1) = P(N_1) + \tfrac{1}{2}P(N_2) = 0.15.$$
$$Q(A_1, [A_{-1}] \mid e_2) = \tfrac{1}{3}P(N_5) = 0.0667.$$
$$Q(A_2, [A_{-2}] \mid e_2) = \tfrac{1}{3}P(N_5) = 0.0667.$$
$$Q(A_3, [A_{-3}] \mid e_2) = \tfrac{1}{2}P(N_3) + P(N_6) = 0.35.$$
$$Q(A_4, [A_{-4}] \mid e_2) = \tfrac{1}{2}P(N_3) = 0.05.$$
$$Q(A_5, [A_{-5}] \mid e_2) = P(N_4) + \tfrac{1}{3}P(N_5) = 0.2667.$$
$$Q(A_6, [A_{-6}] \mid e_2) = P(N_1) + P(N_2) = 0.2.$$
$$Q(A_1, [A_{-1}] \mid e_3) = \tfrac{1}{2}P(N_4) + P(N_5) = 0.3.$$
$$Q(A_2, [A_{-2}] \mid e_3) = 0.$$
$$Q(A_3, [A_{-3}] \mid e_3) = \tfrac{1}{2}P(N_3) + P(N_6) = 0.35.$$
$$Q(A_4, [A_{-4}] \mid e_3) = \tfrac{1}{2}P(N_3) = 0.05.$$
$$Q(A_5, [A_{-5}] \mid e_3) = \tfrac{1}{2}P(N_4) = 0.1.$$
$$Q(A_6, [A_{-6}] \mid e_3) = P(N_1) + P(N_2) = 0.2.$$

Now, since the importance of each expert is given by the weighting vector $(0.2, 0.4, 0.4)$, we can obtain the preference degree of each alternative:

$$\begin{aligned} Q(A_1, [A_{-1}]) &= Q(A_1, [A_{-1}] \mid e_1)0.2 + Q(A_1, [A_{-1}] \mid e_2)0.4 \\ &\quad + Q(A_1, [A_{-1}] \mid e_3)0.4 = 0.4\ \ 0.2 + 0.0667\ 0.4 + 0.3\ 0.4 = 0.22667. \end{aligned}$$

And similarly:

$$Q(A_2, [A_{-2}]) = 0.0667\ \ 0.4 = 0.02667.$$
$$Q(A_3, [A_{-3}]) = 0.3\ \ 0.2 + 0.35\ 0.4 + 0.35\ 0.4 = 0.34.$$
$$Q(A_4, [A_{-4}]) = 0.15\ \ 0.2 + 0.05\ 0.4 + 0.05\ 0.4 = 0.07.$$
$$Q(A_5, [A_{-5}]) = 0.2667\ \ 0.4 + 0.1\ 0.4 = 0.14667.$$
$$Q(A_6, [A_{-6}]) = 0.15\ \ 0.2 + 0.2\ \ 0.4 + 0.2\ 0.4 = 0.19.$$

Thus, general statistical preference gives A_3 as the preferred alternative: $A_3 \ {}_{SP}\ [A_{-3}]$; A_1 is the second preferred alternative, A_6 the third, A_5 the fourth, A_4 the fifth and finally A_2 is the less preferred alternative. Consequently, creating a new product for medium-income customers seems to be the best option, while the worst alternative is creating a new product for high-income customers.

3.5 Conclusions

Sto chastic orders are to ols that allow us to compare random quantities, so they b ecome particularly useful in decision problem s under uncertainty. One of the most imp or- tant sto chastic orders that can be found in the literature is sto chastic dominance. This metho d, based onthe comparison of the cumulative distribution functions, has been widely studied in the literature, and it has b een ap plied in many different areas. One alternative sto chastic order is statistical preference, which has remained unexplored fora long time. For th is reason, we have dedicated the first part of this chapter to the investi- gation of the prop erties of statistical preference as a stochastic order. In particular, while sto chastic dominance is close to the exp ectation, we have seen that statistical preference is related to another lo cation parameter: the median. This showed that both sto chastic orders have adifferent philosophy under their definition.

Interestingly, there are situations where b oth sto chastic orders give rise to the same conclusions. For instance, we have found conditions under which first degree sto chastic dominance implies statistical preference. These situations included, for example, inde- pendent random variables or continuous comonotonic/countermonotonic random vari- ables, among others. Although the two metho ds are not equivalent in general, we have proved that the coincide when comparing indep endent random variables whose distribu- tions are Bernoulli, exp onential, uniform, Pareto, b eta and normal.

Both metho ds have been devised for the pairwise comparison of random variables, and may be unsuitable when more than two random variables must be compared simul- taneously. For this reason, we have intro duced a new stochastic order,that generalises statistical preference and preserves its underlying philosophy, that allows us to comp are more than two random variable s at the same time. We have also investigated its main prop erties and its connection with the usual sto chastic orders.

Sto chastic orders app ears in many different real-life problems. For this reason, the last part of this chapter was devoted to present a numb er of applications that show the relevance of our res ults. On the one hand, we have seen that b oth sto chastic dominance and statistical preference could be an interesting alternative to the comparison of fitness values, andon the other hand we have applied the general statistical preference toa multicriteria decision making with linguistic lab els.

From the results we have showed in this chapter new op en problems arise. For instance, we have given some conditions under which first degree sto chastic dominance implies statistical preference, and we have seen that this relation do es not hold in general. Thus, anaturalquestion arises: is it p ossible to characteris e the situations in which first degree sto chastic dominance implies statistical preference?

Moreover, we have also seen that both sto chastic dominance and statistic prefer- ence coincide for the comparison of indep endent random variables whose distribution is

Bernoulli, exp onential, normal, .. . In fact, b oth metho ds reduce to the comparison of the exp ectation of the variables. We conjecture that for indep endent random variables whose distribution belongs to the exp onential family of distributions, both sto chastic dominance and statistical preference coincide and are equivalent to the comparison of the exp ectation. Although this is an op en question that has not been answered yet, a first approach, basedon simulations, has already be done by Casero ([32]). We have intro duced the general statistical preference as a sto chastic order for the comparison of more than two random variables simultaneously. Although we have investigated its main prop erties, a different approach could be given to this notion. In fact, the gen eral statistical preference could be seen as a fuzzy choice function ([81]) on a set of random variables, since it gives degrees of preferenc e of a random variable over a set a random variables. Then, the investigation of the prop erties of the general statis tical preference as a fuz zy choice function could b e an interesting line of research.

4 Comparison of alternatives under uncertainty and imprecision

In the previous chapter we have dealt with the comparison of alternatives under uncertainty. When these alternatives are modelled by means of random variables, the comparison must be performed using stochastic orders. However, there are situations in which it is not possible or adequate to model the experiments by means of a single random variable, due to the presence of imprecision in the experiment. In other words, we focus now in situations where the alternatives are defined under uncertainty but also under imprecision. In such cases, we shall compare sets of random variables instead of single ones; more generally, we shall compare imprecise probability models. For this reason, this chapter is devoted to the extension of the pairwise methods studied in the previous chapter to the comparison of imprecise probability models.

As we have already mentioned, imprecise probabilities ([205]) is a generic term that refers to all mathematical models that serve as an alternative and a generalisation to probability models in case of imprecise knowledge. In this respect, stochastic dominance was connected to imprecise probabilities by Denoeux ([61]), who generalised this notion to the comparison of belief functions ([187]). He proposed four extensions of stochastic dominance based on the orders between real intervals given in [78]. One step forward was made by Aiche and Dubois ([1]), by using stochastic dominance to compare random intervals stemming from rankings between real intervals, in a similar manner as Denoeux, and also in the comparison of fuzzy random variables ([105]).

On the other hand, the comparison of sets of random variables appears naturally in decision making under imprecision. In this sense, the usual *utility order* has already been extended in several ways to the comparison of sets of random variables: interval dominance ([219]), maximax ([184]) and maximin criteria ([82]), and E-admissibility ([107]). See a survey on this topic in ([202]).

With respect to statistical preference, Couso and Sánchez ([46]) proposed it as a method for comparing sets of desirable gambles (see [205, Sec. 2.2.4] for further information). Also, Couso and Dubois ([43]) proposed a common formulation for both statistical preference and stochastic dominance to the comparison of imprecise probability models, and they studied its formulation in terms of expected utility.

Our aim here is to consider a more general situation: we start from a binary relation, that may be stochastic dominance, statistical preference or any other, as in Section 2.1, and extend it to the comparison of sets of random variables. We shall consider six possible extensions of the binary relation, and we shall study the connections between them. Afterwards, we consider the particular cases when the binary relation is stochastic dominance or statistical preference. As we shall see, our approach is more general than that of Denoeux, since the comparison of belief functions arises a particular case. On the other hand, our approach differs from the one of [43, 46] because they considered the comparison of sets of desirable gambles instead of sets of random variables, and the underlying philosophy of their approach is slightly different to ours.

After these general considerations, we shall focus on two scenarios that can be embedded into the comparison of sets of random variables: the comparison of two alternatives with imprecision either in the utilities or in the beliefs. The former will be formulated by means of random sets, and their comparison will be made by means of the associated sets of measurable selections. In the latter, we shall assume that there is a set of probability measures modelling the real probability measure of the probability space.

Since there could be imprecision on the initial probability, we devote the next section to the modelling of the joint distribution in an imprecise framework. For this aim, we shall investigate how the bivariate distribution can be expressed when there is imprecision in the initial probability. Then, we investigate bivariate p-boxes, and in particular how sets of bivariate distribution functions can define a bivariate p-box, and we study if it is possible to formulate an imprecise version of the famous Sklar's Theorem (see Theorem 2.27).

We conclude the chapter with several applications. First of all, we use imprecise stochastic dominance to compare sets of Lorenz Curves and cancer survival rates. Secondly, we use a multicriteria decision making problem to illustrate how imprecise stochastic orders can be applied in a context of imprecision either in the utilities or in the beliefs.

4.1 generalisation of the binary relations to the comparison of sets of random variables

In the following, we propose a number of methods for comparing pairs of sets of variables which are based on performing pairwise comparisons of elements within these sets. First we shall give our definitions for the case where the comparisons of the elements are made by means of a binary relation, as we did at the beginning of Section 2.1, and later we shall apply them to the particular cases where this binary relation consists of sto chastic dominanceor statistical preference.

We shall consider a probability space $(\Omega, \mathcal{A}, P)$ and an ordered utility scale Ω, that in some situations will b e consid ered as numerical. We shall also conside r sets of random variables, defined fromthe probability space to Ω, that will be denoted by $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \ldots$.

We begin with the extension of a binary relation to the comparison of sets of random variables.

Definition 4.1 *Let be a binary relation betweenrandom variables defined froma probability space $(\Omega, \mathcal{A}, P)$ to an ordered utility scale Ω. Given twosetsof random variables $\mathcal{X}$ and $\mathcal{Y}$, we say that:*

1. *$\mathcal{X}$ $_1$ $\mathcal{Y}$ if andonly if forevery X $\mathcal{X}$, Y $\mathcal{Y}$ it holdsthat X Y.*
2. *$\mathcal{X}$ $_2$ $\mathcal{Y}$ if andonly if thereis some X $\mathcal{X}$ such that X Y for every Y $\mathcal{Y}$.*
3. *$\mathcal{X}$ $_3$ $\mathcal{Y}$ if andonly if forevery Y $\mathcal{Y}$ there issome X $\mathcal{X}$ such that X Y.*
4. *$\mathcal{X}$ $_4$ $\mathcal{Y}$ if andonly if thereare X $\mathcal{X}$, Y $\mathcal{Y}$ such that X Y.*
5. *$\mathcal{X}$ $_5$ $\mathcal{Y}$ if andonly if thereis some Y $\mathcal{Y}$ such that X Y for every X $\mathcal{X}$.*
6. *$\mathcal{X}$ $_6$ $\mathcal{Y}$ if andonly if forevery X $\mathcal{X}$ there is Y $\mathcal{Y}$ such that X Y.*

Remark 4.2 *As wedid inDefinition 2.1, from any of these definitions we can infer immediately arelation of* strict *preference($_i$) and the* indifference*($\equiv_i$):*

$$\mathcal{X} \ _i \ \mathcal{Y} \quad \mathcal{X} \ _i \ \mathcal{Y} \text{ and } \mathcal{Y} \ _i \ \mathcal{X},$$
$$\mathcal{X} \equiv_i \mathcal{Y} \quad \mathcal{X} \ _i \ \mathcal{Y} \text{ and } \mathcal{Y} \ _i \ \mathcal{X},$$

for any $i = 1, \ldots, 6$. Moreover, wesaythat $\mathcal{X}$ and $\mathcal{Y}$ are incomparablewithrespectto $_i$ when $\mathcal{X}$ $_i$ $\mathcal{Y}$ and $\mathcal{Y}$ $_i$ $\mathcal{X}$.

The conditions in this definition can be given the following interpretation. $_1$ means that any alternative in $\mathcal{X}$ is -preferredto anyalternative in $\mathcal{Y}$, and as such it is related to the idea of interval dominance from decision making with sets of probabilities [219]. Conditions $_2$ and $_3$ mean that the "b est" alternative in $\mathcal{X}$ is -better than the "best" alternative in $\mathcal{Y}$. The difference b etween th em lies in whether there is a maximal element in $\mathcal{X}$ in theorderdetermined by . Thesetwoconditionsare relatedtothe Γ-maximax criteria considered in [184]. Ontheother hand, conditions $_5$ and $_6$ mean that the "worst" alternative in $\mathcal{X}$ is -preferredto the"worst" alternative in $\mathcal{Y}$, and arerelated to the Γ-maximin criteria in [20, 82]. Again, the difference between them lies in whether there is a minimum element in $\mathcal{Y}$ with resp ect to the order determined by or not.

Finally, $_4$ is a weakenedversion of $_1$, in thesense thatit onlyrequiresthatsome alternative in X is -preferred tosome otheralternativein Y, instead of requiring it for any pair in X, Y.

Taking this interpretation into account, it is not difficult to establish the followi ng relationships b etween the definitions.

Prop osition 4.3 *The fol lowing implications hold:*

- $_1$ $_2$ $_3$ $_4$.
- $_1$ $_5$ $_6$ $_4$.

Pro of ($_1$ $_2$): If X Y for every X X , Y Y , in particular given any X X it holds that X Y for every Y Y .

($_2$ $_3$): If there exists X X such that X Y for every Y Y , the condition in $_3$ is satisfied with resp ect to X for every Y Y .

($_3$ $_4$): Iffor every Y Y there exists X_Y X such that X_Y Y, we havea pair (X_Y, Y) $X \times Y$ such that X_Y Y.

($_1$ $_5$): If X Y for every X X and Y Y , in particular given any Y Y it holds that X Y for every X X .

($_5$ $_6$): If there is some Y Y such that X Y for every X X , in particul ar, for every X X itholds that X Y.

($_6$ $_4$): If for every X X there exists Y_X Y such that X Y_X , we havea pair (X, Y_X) $X \times Y$ such that X Y_X . ■

The previous impli cations are depicted in Figure 4.1. Oth er relationships b etween the six definitions do not hold in general, as we can see in the following example.

Example 4.4 *Considera probability space withonly oneelement ω, and let δ_x denote the random variable satisfy ing $\delta_x(\omega) = x$. Consideralsothe binaryrelation such that:*

$$X \quad Y \quad X(\omega) \geq Y(\omega). \tag{4.1}$$

If we take $X = \{\delta_1, \delta_3\}$ and $Y = \{\delta_2\}$, it fol lows that δ_3 δ_2 δ_1, whence, applying Definition 4.1, we have that:

$$X \ _2 Y, \quad X \ _3 Y, \quad X \equiv_4 Y, \quad Y \ _5 X, \quad Y \ _6 X$$

and X and Y areincomparablewith respectto thefirst extension. From thiswe deduce that:

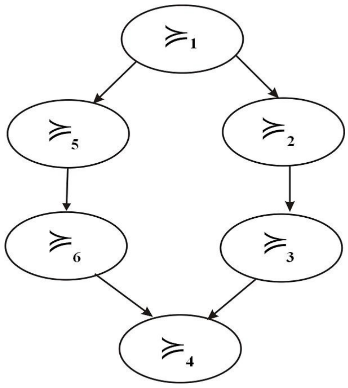


Figure 4.1: Relationships among the diffe rent extensions of the binary relation for the comparison of setsof random variables.

- 2 , 1, 5, 6 *and therefore* 3 , 1, 5, 6.
- 4 , 1, 2, 3, 5, 6.
- 5 , 1, 2, 3 *and therefore* 6 , 1, 2, 3.

Next, given $X = Y = \{\delta_x : x \quad (0,1)\}$, *we have that* $X \equiv_3 Y$ *and* $X \equiv_6 Y$, *because* $\delta_x \equiv \delta_x$ *for all* $x \quad (0,1)$. *However,* X *and* Y *areincomparablewith respect tosecond andfifth definitions, because there are not* $x_1, x_2 \quad (0,1)$ *for which* $\delta_{x_1} \quad \delta_r$ *and* $\delta_r \quad \delta_{x_2}$ *for all* $r \quad (0,1)$. *Hence:*

- 3 2.
- 6 5.

Remark 4.5 *Insomecases, itmaybe interesting tocombinesomeof thesedefinitions, for instance to consider* X *preferred to* Y *when it ispreferred according to definitions* $_2$ *and* $_5$. *Takinginto account theimplications depictedin Proposition 4.3, the combinations thatproduce newconditions arethose where we take onecondition out of* $\{\ _2, \ _3\}$ *together with one out of* $\{\ _5, \ _6\}$.

If we combine for instance $_2$ *with* $_5$, *we can introduce the extension, denoted by* $_{2,5}$, *and defined by:*

$$X \quad _{2,5} Y \quad X \quad _2 Y \text{ and } X \quad _5 Y.$$

Then, $_{2,5}$ *requires that* X *hasa -bestcasescenario which isbetterthan anysituation in* Y *and that* Y *hasa -worstcase whichisworse thanany situationin* X. *This turns* *out to be an intermediate condition bet ween $_1$ and each of $_2$ and $_5$, anditcan be derivedfrom the previousexample that itis not equivalentto any of them.*

The implications in Prop osition 4.3 can also be seen easily in the case where X and Y are finite sets, $X = \{X_1, \ldots, X_n\}$ and $Y = \{Y_1, \ldots, Y_m\}$. Then ifwedenote by M the $n \times m$ matrix where

$$M_{i,j} = \begin{cases} 1 & \text{if } X_i \quad Y_j \\ 0 & \text{otherwise} \end{cases}$$

the ab ove definitions are characterised in the following way:

- $X \quad _1 Y \quad M = 1_{n,m}$.
- $X \quad _2 Y \quad i \quad \{1, \ldots, n\}$ such that $M_{i,} = 1_{1,m}$.
- $X \quad _3 Y \quad j \quad \{1, \ldots, m\}$ such that $M_{,j} = 0_{n,1}$.
- $X \quad _4 Y \quad M = 0_{n,m}$.
- $X \quad _5 Y \quad j \quad \{1, \ldots, m\}$ such that $M_{,j} = 1_{n,1}$.
- $X \quad _6 Y \quad i \quad \{1, \ldots, n\}$ such that $M_{i,} = 0_{1,m}$.

Observe that, aswe havealready seen, forany binaryrelation , its extensions $_2$ and $_3$ (resp ectively $_5$ and $_6$) are quite related: b oth compare the b est (resp ectively, the worst) alternatives withineach set X, Y. Since the difference between them lies on whether there is a maximal (resp ectively, minimal) element within each of these sets or not, we can easily give a necess ary and sufficient condition for the equivalences $_2 \quad _3$ and $_5 \quad _6$.

Prop osition 4.6 *Let be a binaryrelation on the set of random variables that is reflexive and transitive.*

(a) Given aset X of random variables, $X \quad _3 Y \quad X \quad _2 Y$ foranyset of variables Y if and only if X hasa maximumelement under .

(b) Given a set Y of random variables, $X \quad _6 Y \quad X \quad _5 Y$ foranyset of variables X if and only if Y has a minimumelementunder .

Pro of

(a) Assume that X has a maximumelement X such that $X \quad X$ for every $X \quad X$. If $X \quad _3 Y$, then for every $Y \quad Y$ there is some $X_Y \quad X$ such that $X_Y \quad Y$. Since is transitive, wededuce that X X_Y Y, and then X Y for every Y Y, and as aconsequence X $_2$ Y.

Conve rs ely, if X do es not have a maximum element,we can take Y = X and we would have $X \equiv_3 Y$ because is reflexive; however, X and Y are incomparable with resp ect to $_2$ because X do es not have a maximum element.

(b) Similarly, if Y hasa minimum element Y, it holds that Y Y for any Y Y. If X $_6$ Y, then for every X X there exists Y_X Y such that X Y_X, and since is transitive we obtain that X Y for every X X, whence X $_5$ Y.

Conve rs ely, if Y do es not have a minimum element, wecan take X = Y and we would have $X \equiv_6 Y$ because is reflexive; however, X and Y are incomparable with resp ect to $_5$ because Y do es not have a minimum element. ■

Under some conditions, we can also give a simpler characterisation of the ab ove prop erties:

Prop osition 4.7 *Let bea binaryrelationbetweenrandomvariables, andassumethat it satisfies the* Pareto Dominanc e *condition:*

$$X(\omega) \geq Y(\omega) \quad \omega \quad X \quad Y. \tag{4.2}$$

Considertwo sets of random variables X, Y. Ifthe randomvariables $\min X$, $\max X$ *exist and belong to X and* $\min Y$, $\max Y$ *exist andbelong to Y, then:*

(a) X $_1$ Y $\min X$ $\max Y$.

(b) X $_2$ Y X $_3$ Y $\max X$ $\max Y$.

(c) X $_4$ Y $\max X$ $\min Y$.

(d) X $_5$ Y X $_6$ Y $\min X$ $\min Y$.

Pro of Note thatwhenboth X, Y includea maximum anda minimumrandom variable, Equation (4.2) implie s that for every X X, Y Y,

$$\min X \quad Y \quad X \quad Y \quad \max X \quad Y$$

and

$$X \quad \max Y \quad X \quad Y \quad X \quad \min Y.$$

Then:

(a) Since $\min X$ $\max Y$, it isobviousthat X $_1$ Y. Ontheother hand, using the pre vious equations, if every X X and Y Y satisfy X Y, then also $\min X \geq \max Y$.

(c) Since max$\mathcal{X}$ min $\mathcal{Y}$, and max$\mathcal{X}$ X and min $\mathcal{Y}$ Y, then $\mathcal{X}$ $_4$ $\mathcal{Y}$. On the other hand, using the previous e quations,if X Y for some X $\mathcal{X}$, Y $\mathcal{Y}$, also max$\mathcal{X}$ min $\mathcal{Y}$.

(b,d) Using the previousequations, $\mathcal{X}$ has a maximumelement and $\mathcal{Y}$ has a min imum element under . By Prop osition 4.6, $\mathcal{X}$ $_3$ $\mathcal{Y}$ $\mathcal{X}$ $_2$ $\mathcal{Y}$ and $\mathcal{X}$ $_6$ $\mathcal{Y}$ $\mathcal{X}$ $_5$ $\mathcal{Y}$. The remaining equivalence can b e estab lished in an analogous manner to the previous cases. ■

Remark 4.8 *AccordingtoRemark4.5, underthe conditions ofthepreviousresult, itis immediate that* $\mathcal{X}$ $_{2,5}$ $\mathcal{Y}$ *if andonly if* max$\mathcal{X}$ max$\mathcal{Y}$ *and* min$\mathcal{X}$ min $\mathcal{Y}$.

Next we investigate which prop erties of the binary relation hold ontothe extensions $_1, \ldots,$ $_6$. Obviously, since all these definitions become inthe caseof sin gle tons, if is notreflexive (resp., antisymmetric, transitive), neither are $_i$, for $i = 1, \ldots, 6$. Converse ly, we can establish the following result.

Prop osition 4.9 *Let bea binaryrelationonrandomvariables, andlet* $_i$, $i = 1, \ldots, 6$ *be its extensions to sets of random variables, given by Definition 4.1.*

(a) If is reflexive,soare $_3$, $_4$ *and* $_6$.

(b) If is antisymmetric,so is $_1$.

(c) If is transitive,so are $_i$ *for* $i = 1, 2, 3, 5, 6$.

Pro of Firstofall, if is reflexive, $X \equiv X$ for anyrandom variable X, and applying Definition 4.1 we de duce that$\mathcal{X}$ $_i$ $\mathcal{X}$ for any $i = 3, 4, 6$ and any set of random variables $\mathcal{X}$.

Secondly, assume that is antisymmetric andthat two sets of random variables $\mathcal{X}, \mathcal{Y}$ satisfy $\mathcal{X}$ $_1$ $\mathcal{Y}$ and $\mathcal{Y}$ $_1$ $\mathcal{X}$. Then, X Y and Y X for every X $\mathcal{X}$ and Y $\mathcal{Y}$, and by the antisymmetry prop erty of , wededuce that $X = Y$ for every X $\mathcal{X}$, Y $\mathcal{Y}$. But this can only be if $\mathcal{X} = \{Z\} = \mathcal{Y}$ forsome randomvariable Z. As a consequence, $_1$ is antisymmetric.

Finally, assume that istransitive, andletusshowthatsoare $_i$ for $i = 1, 2, 3, 5, 6$. Considerthree setsof random variables $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$:

1. If $\mathcal{X}$ $_1$ $\mathcal{Y}$ and $\mathcal{Y}$ $_1$ $\mathcal{Z}$ then X Y and Y Z for every X $\mathcal{X}$, Y $\mathcal{Y}$, Z $\mathcal{Z}$. Applying the transitivity of , we dedu ce that X Z for every X $\mathcal{X}$, Z $\mathcal{Z}$, and as aconsequence $\mathcal{X}$ $_1$ $\mathcal{Z}$.

2. If $X \quad _2 Y$ and $Y \quad _2 Z$, there is $X \quad X$ such that $X \quad Y$ for every $Y \quad Y$ and there is $Y \quad Y$ such that $Y \quad Z$ for every $Z \quad Z$. In particular, $X \quad Y \quad Z$ for every $Z \quad Z$, whence, by the trans itivity of , $X \quad _2 Z$.

3. If $X \quad _3 Y$ and $Y \quad _3 Z$, for every $Y \quad Y$ thereis some $X_Y \quad X$ such that $X_Y \quad Y$, and for every $Z \quad Z$ there is $Y_Z \quad Y$ such that $Y_Z \quad Z$. As a consequence,for every $Z \quad Z$ itholds that $X_{Y_Z} \quad Z$, and th erefore $X \quad _3 Z$.

The pro of of the transitivity of $_5$ and $_6$ holdsbyanalogyto thatof $_2$ and $_3$, resp ectively. ∎

Our next example shows that reflexivity an d antisymmetry do not hold for definitions different than the ones of s tate ments (a) and (b).Toshow thatthe fourthextension is not transitive in general, even whenthe binaryrelationis, werefer toExample4.18, where we shall show that the fourth extension is not tran sitive when considering the binary relation to be the first degree stochastic dominance.

Example 4.10 *Consider theuniverse* $\Omega = \{\omega\}$ *and, as we made inExample 4.4, denote by* δ_x *therandomvariable suchthat* $\delta_x(\omega) = x$, *andthebinary relationdefinedin Equation* (4.1). *Considerthe set ofrandom variables* X *defined by* $X = \{\delta_x : x \quad (0, 1)\}$. *Then, although is reflexive,* X *is incomparable wit h itself with respect to* $_1$, $_2$ *and* $_5$. *Now,consider the setsof random variables* X *and* Y *defined by:*

$$X = \{\delta_x : x \quad [0, 1]\} \text{ and } Y = \{\delta_x : x \quad [0, 1] \{0.5\}\}.$$

Then, $X \equiv_i Y$ *for any* $i = 2, 3, 4, 5, 6$, *but* $X = Y$, *while is an antisymmetricrelation.*

Another interesting prop erty in a binary relation is that of *completeness*, which means thatgiven anytwoelements, either oneis preferredtothe otherorthey are indifferent, but theyare never incomparable. From Prop osition 4.3, it follows that the incomparable pairs with resp ect to an extension $_i$ are also incomparable with resp ect to the stronger extensions. The following resultshowsthat if is a complete relation, then itsweakest extensions (namely, $_3$, $_4$ and $_6$) also induce complete binary relations:

Prop osition 4.11 *Considera binary relation betweenrandomvariables, andlet* $_i$, *for* $i = 1, \ldots, 6$, *be its extensions to sets of random variables given by Definition 4.1. If is complete, then so are* $_3$, $_4$ *and* $_6$.

Pro of Let X, Y be two sets of random variables, and assumethat $X \quad _3 Y$. Then there is some $Y \quad Y$ such that $X \quad Y$ for all $X \quad X$. But since is acomplete relation, this meansthat $Y \quad X$ for all $X \quad X$. As a consequence, $Y \quad _2 X$, and applying Prop osition 4.3 we deduce that $Y \quad _3 X$. Hence , the binary relation $_3$ is complete.

	1	2	3	4	5	6
Reflexive			•	•		•
Antisymmetric	•					
Transitive	•	•	•		•	•
Complete			•	•		•

Table 4.1: Summary of the prop erties of the binary relation that hold onto their extensions $_1, \ldots, _6$.

Onthe otherhand, if $X \; _4 \; Y$, we deduce from Prop osition 4.3 that also $X \; _3 \; Y$, whence the ab ove reasoning implies that $Y \; _3 \; X$ and again from Prop osition 4.3 we deduce that $Y \; _4 \; X$.

The pro of that $_6$ alsoinduces a completerelation isanalogous. ∎

Letus now give an example where we see that the completenessof the binary relationship do es not imply the completeness of the extensions $_1, \; _2, \; _5$.

Example 4.12 *Consideragain Example4.10, andtake thesets of random variables $X = Y = \{\delta_x : x \;\; (0,1)\}$ and the binary relation definedinEquation* (4.1). *Although is complete, X and Y are incomparable with respectto $_1, \; _2$ and $_5$.*

Table 4.1 summarises the properties we have investigated in Prop ositions 4.9 and 4.11.

Remark 4.13 *Althoughin this report we shall focus on the particular application of Definition4.1 to the relation associated with stochastic dominance or statisticalpref- erence, there are other cases of interest. Perhapsthemost important oneis thatwhere the comparison between pairs of random variables is made by means of their expected utility:*

$$X \quad Y \quad E(X) \geq E(Y);$$

it is not difficult to see that Definition 4.1 gives rise to some wel l-known generalisations of expectedutility thatare formulatedinterms of lowerandupper expectations. Consider two sets X, Y andassume thatthe expectationsof all their elements exist. Then with respect to definition $_1$ it holdsthat:

$$X \;\; _1 \; Y \quad \underline{E}(X) = \inf_{X \; X} E(X) \geq \sup_{Y \; Y} E(Y) = \overline{E}(Y),$$

which relates this notion tothe concept of intervaldominance *considered in [219].*

If we now consider definition $_3$, it holds that

$$X \;\; _3 \; Y \quad \overline{E}(X) = \sup_{X \; X} E(X) \geq \sup_{Y \; Y} E(Y) = \overline{E}(Y).$$

Thus, definition $_3$ *is stronger than the* maximax*criterium [184], whichis based on comparing the best possibilities in our sets of alternat ives. Similarly, if we consider definition* $_6$ *it holdsthat:*

$$X \quad _6\, Y \qquad \underline{E}(X) = \inf_{X \in \mathcal{X}} E(X) \geq \inf_{Y \in \mathcal{Y}} E(Y) = \underline{E}(Y).$$

Thus, definition $_6$ *isstrongerthanthe* maximin *criterium[82], whichcompares the worst possibilities within the sets of alternatives.*

Final ly, definition $_4$ *implies that*

$$X \quad _4\, Y \qquad \overline{E}(X) = \sup_{X \in \mathcal{X}} E(X) \geq \inf_{Y \in \mathcal{Y}} E(Y) = \underline{E}(Y),$$

so if X *is* $_4$*-preferred to* Y *then it is also preferred with respect to the criterion of* E-admissibility *from [107]. See [43, 202]for relatedcomments.*

4.1.1 Imprecise sto chastic dominance

Inthis subsection, weexplore insome detailthe casewherethebinaryrelation is the one asso ciated with the notion of first degree sto chastic dominance we have intro duced inDefinition2.2, i.e., the relation is defined by $_{\text{FSD}}$. We call this extension impre cise sto chastic dominance. We shall assume that the utility space Ω is $[0, 1]$ although the results can be immediately extended to any bounded interval of real numb ers. Since sto chastic dominance is based on the comparison of cumulative distribution functions asso ciated with the random variables, we shall employ the notation $\mathcal{F}_X \;_{\text{FSD}}\; \mathcal{F}_Y$ instead of $\mathcal{X} \;_{\text{FSD}}\; \mathcal{Y}$. For the same reason, along this subs ection we will consider sets of cumulative distribution functions $\mathcal{F}_X$ and $\mathcal{F}_Y$ insteadofsetsofrandom variables $\mathcal{X}$ and $\mathcal{Y}$.

Remark 4.14 *Fromnow on, we shall say that aset of distribution functions* $\mathcal{F}_X$ *is* (FSD_i)*-preferred or that it* (FSD_i)*-stochastical ly dominates another set of distribution functions* $\mathcal{F}_Y$ *when* $\mathcal{F}_X \;_{\text{FSD}_i}\; \mathcal{F}_Y$. *We wil l also use the notation* $_{\text{FSD}_{i,j}}$ *when both* $_{\text{FSD}_i}$ *and* $_{\text{FSD}_j}$ *hold.*

An illustration of the six extensions of Definition 4.1 when considering sto chastic dominance isgiven in Figure4.2, where we compare the set of distribution functions re presented by a continuous line(that weshall call continuous distributions inthis paragraph) with the set of distribution functions represented by a dotte d line (that we shall call dotted distributions). Ontheone hand, in the left picture the set of continuous distributions (FSD1)-sto chastically dominates the set of dotted distributions. In theright picture, there is a continuous distribution that dominates all dotted distributions, and a dotted distribution whichis dominated by all continuous distributions. T his means thatthe set of continuous distributions sto chastically dominates the set of dotted distributions with resp ect to the second to sixth definitions. Sincethere isalsoadotted distribution that is dominated by a continuous distribution, we deduce that the set of continuous distributions and the set of dotted distributions are equivalent with resp ect to the fourth definition. Noticethat the binaryrelationship considered inExample 4.4

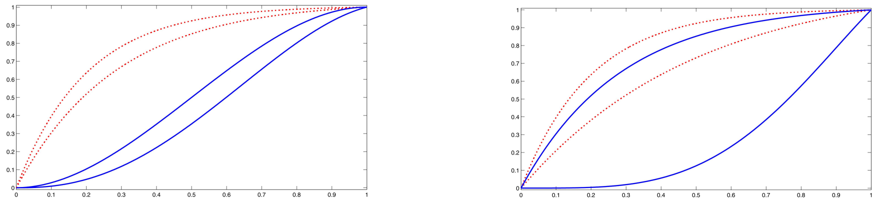


Figure 4.2: Examples of several definitions of imprecise sto chastic dominance.

is equivalent to first degree sto chastic dominance when the initial space Ω only has one element. Then, such example shows that the converse implications of Prop osition 4.3 do not hold in general when considering the binary relation to be the first degree sto chastic dominance.

Now, we investigate which prop erties hold when considering the strict imprecise sto chastic dominance.

Prop osition 4.15 *Consider the extensions of stochastic dominance given in Definition 4.1. Itholds that:*

- $F_X \quad {}_{FSD_2} F_Y \quad F_X \quad {}_{FSD_3} F_Y$.
- $F_X \quad {}_{FSD_5} F_Y \quad F_X \quad {}_{FSD_6} F_Y$.

Pro of Webegin proving that ${}_{FSD_2}$ implies ${}_{FSD_3}$. Observe that $F_X \quad {}_{FSD_2} F_Y$ is equivalent to:

$$(I) \quad F_X \quad {}_{FSD_2} F_Y \qquad F_1 \quad \mathcal{F}_X \text{ such that } F_1 \leq F_2 \text{ for all } F_2 \quad \mathcal{F}_Y.$$
$$(II) \quad F_Y \quad {}_{FSD_2} F_X \qquad F_2 \quad \mathcal{F}_Y, \ F_1 \quad \mathcal{F}_X \text{ such that } F_2 \leq F_1.$$

It follows from (I) and Prop osition 4.3 that $F_X \quad {}_{FSD_3} F_Y$. We only have to prove that $F_Y \quad {}_{FSD_3} F_X$, or equivalently, that there is $F_1 \quad \mathcal{F}_X$ such that $F_2 \leq F_1$ for any $F_2 \quad \mathcal{F}_Y$. If F_1 satisfies this prop erty, the pro of is finished. If not, there issome $F_2 \quad \mathcal{F}_Y$ such that $F_2 \leq F_1$, wh ence $F_1 = F_2$. Applying (II), thereexists some $F_1 \quad \mathcal{F}_X$ such that $F_1 \leq F_1$, which meansthat $F_1(t) < F_1(t)$ for some t. As a consequence, $F_1(t) < F_2(t)$ for any $F_2 \quad \mathcal{F}_Y$, whence $F_Y \quad {}_{FSD_3} F_X$. Hence, $F_X \quad {}_{FSD_3} F_Y$.

Let us now prove that $_{FSD_5}$ $_{FSD_6}$. Similarly to the previous case, F_X $_{FSD_5}$ F_Y is equivalent to:

$$(I) \quad F_X \;_{FSD_5}\; F_Y \qquad F_2 \quad F_Y \text{ such that } F_1 \leq F_2 \text{ for all } F_1 \quad F_X.$$
$$(II) \quad F_Y \;_{FSD_5}\; F_X \qquad F_1 \quad F_X,\; F_2 \quad F_Y \text{ such that } F_2 \leq F_1.$$

It follows from (I) and Proposition 4.3 that F_X $_{FSD_6}$ F_Y. We only have to prove that F_Y $_{FSD_6}$ F_X, or equivalently, that there is F_2 F_Y such that $F_2 \leq F_1$ for any F_1 F_X. If F_2 satisfies this property, the proof is finished. If not, there exists F_1 F_X such that $F_2 \leq F_1$, and applying (I) we deduce that $F_1 = F_2$ F_X. Applying (II) we deduce that there is some F_2 F_Y such that $F_2 \leq F_1$, whence there is some t such that $F_2(t) > F_1(t) = F_2(t) \geq F_1(t)$ for every F_1 F_X. Hence, $F_2 \leq F_1$ for any F_1 F_X and the property holds. ∎

Furthermore, next example shows that there are no other relationships between the strict extensions of stochastic dominance.

Example 4.16 *Consider the same conditions of Example 4.4:* $\Omega = \{\omega\}$, δ_x *is the random variable given by* $\delta_x(\omega) = x$ *and* *is given by Equation* (4.1), *that is equivalent to* $_{FSD}$ *in this case.*

Take the sets $X = \{\delta_1\}$ *and* $Y = \{\delta_0, \delta_1\}$. *It holds that:*

$$X \;_{FSD_1}\; Y \text{ and } X \;_{FSD_6}\; Y,$$

but $X \equiv_{FSD_2} Y$ *and* $X \equiv_{FSD_4} Y$. *Then,* $_{FSD_1}$ $_{FSD_2}$ *and* $_{FSD_6}$ $_{FSD_4}$.

If we consider the sets $X = \{\delta_0, \delta_1\}$ *and* $Y = \{\delta_0\}$, *it holds that:*

$$X \;_{FSD_1}\; Y \text{ and } X \;_{FSD_3}\; Y,$$

but $X \equiv_{FSD_5} Y$ *and* $X \equiv_{FSD_4} Y$. *Then,* $_{FSD_1}$ $_{FSD_5}$ *and* $_{FSD_3}$ $_{FSD_4}$.

With respect to the other results, since $_{FSD}$ is reflexive and transitive, we can apply Proposition 4.6 and characterise the equivalences between $_{FSD_2}$ and $_{FSD_3}$, and also between $_{FSD_5}$ and $_{FSD_6}$ by means of the existence of a maximum and a minimum value in the sets F_X, F_Y we want to compare. Moreover, we can deduce from Proposition 4.9 and Examples 4.10 and 4.12 that $_{FSD_i}$ is reflexive for $i = 3, 4, 6$ and transitive for $i = 1, 2, 3, 5, 6$. On the other hand, since two different random variables may induce the same distribution function, $_{FSD}$ is not antisymmetric. Nevertheless, if we are dealing with sets of cumulative distribution functions instead of sets of random variables, $_{FSD}$ becomes antisymmetric. Next example shows that (FSD_4) is not transitive in general.

Remark 4.17 *Through this subsection we shall present several examples showing that the propositions established cannot be improved, in the sense that the missing implications* *do not hold in general. Some of these examples will consider distribution functions associated with probability measures with finite supports. To fix notation, given* $a = (a_1, \ldots, a_n)$ *such that* $a_1 + \ldots + a_n = 1$, *and* $t = (t_1, \ldots, t_n)$ *with* $t_1 \leq \ldots \leq t_n$, *the function* $F_{a,t}$ *corresponds to the cumulative distribution function of the probability measure* $P_{a,t}$ *satisfying* $P_{a,t}(\{t_i\}) = a_i$ *for* $i = 1, \ldots, n$. *Indeed, the only continuous distribution function we shall consider is the identity* $F = id$, *defined by* $F(x) = id(x) = x$ *for any* x $[0, 1]$

Example 4.18 *Consider the three sets of cumulative distribution functions* F_X, F_Y *and* F_Z *defined by:*

$$F_X = \{F_{(0.5,0.5),(0,1)}\}, \quad F_Z = \{F\}, \quad F_Y = F_X \quad F_Z.$$

Since both sets F_X *and* F_Z *are included in* F_Y, *Proposition 4.29 later on assures that* $F_X \equiv_{FSD_4} F_Y$ *and* $F_Y \equiv_{FSD_4} F_Z$. *However,* F_X *and* F_Z *are not comparable, since the distribution functions* $F_{(0.5,0.5),(0,1)}$ *and* F *are not comparable with respect to first degree stochastic dominance.*

Since FSD also complies with Pareto dominance (Equation (4.2)), we deduce from Proposition 4.7 that when the sets F_X and F_Y to compare have both a maximum and a minimum element, we can easily characterise the conditions $_{FSD_i}$, $i = 1, \ldots, 6$ by comparing these maximum and minimum elements only. Finally, note that, as we already mentioned in Example 2.3, FSD is not a complete relation, and as a consequence, Proposition 4.11 is not applicable in this context.

As we remarked in Section 2.2.1, p-boxes are one model within the theory of imprecise probabilities. Stochastic dominance between sets of probabilities or cumulative distribution functions can be studied by means of a p-box representation. Given any set of cumulative distribution functions F, it induces a p-box $(\underline{F}, \overline{F})$, as we saw in Equation (2.16):

$$\underline{F}(x) := \inf_{F \quad F} F(x), \quad \overline{F}(x) := \sup_{F \quad F} F(x).$$

Our next result relates the imprecise stochastic dominance for sets of cumulative distribution functions to their associated p-box representation.

Proposition 4.19 *Let* F_X *and* F_Y *be two sets of cumulative distribution functions, and denote by* $(\underline{F}_X, \overline{F}_X)$ *and* $(\underline{F}_Y, \overline{F}_Y)$ *the p-boxes they induce by means of Equation* (2.16). *Then the following statements hold:*

1. $F_X \quad_{FSD_1} F_Y \quad \overline{F}_X \quad_{FSD} \underline{F}_Y$.
2. $F_X \quad_{FSD_2} F_Y \quad \underline{F}_X \quad_{FSD} \underline{F}_Y$.
3. $F_X \quad_{FSD_3} F_Y \quad \underline{F}_X \quad_{FSD} \underline{F}_Y$.

4. $\mathcal{F}_X \;\; _{FSD_4}\; \mathcal{F}_Y \qquad \underline{F}_X \;\; _{FSD}\; \overline{F}_Y$.
5. $\mathcal{F}_X \;\; _{FSD_5}\; \mathcal{F}_Y \qquad \overline{F}_X \;\; _{FSD}\; \overline{F}_Y$.
6. $\mathcal{F}_X \;\; _{FSD_6}\; \mathcal{F}_Y \qquad \overline{F}_X \;\; _{FSD}\; \overline{F}_Y$.

Pro of

(1) Note that $\mathcal{F}_X \;\; _{FSD_1}\; \mathcal{F}_Y$ ifand only if $F_1 \leq F_2$ for every $F_1 \;\; \mathcal{F}_X, F_2 \;\; \mathcal{F}_Y$, and this is equivalent to $\overline{F}_X = \sup_{F_1 \; \mathcal{F}_X} F_1 \leq \inf_{F_2 \; \mathcal{F}_Y} F_2 = \underline{F}_Y$.

(3) By hyp othesis, for every $F_2 \;\; \mathcal{F}_Y$ thereis some $F_1 \;\; \mathcal{F}_X$ such that $F_1 \leq F_2$. Asa consequence, $\underline{F}_X \leq F_2 \;\; F_2 \;\; \mathcal{F}_2 \qquad \underline{F}_X \leq \inf_{F_2 \; \mathcal{F}_Y} F_2 = \underline{F}_Y$.

(4) Ifthere are $F_1 \;\; \mathcal{F}_X$ and $F_2 \;\; \mathcal{F}_Y$ such that $F_1 \leq F_2$, then $\underline{F}_X \leq F_1 \leq F_2 \leq \overline{F}_Y$.

(6) Iffor every $F_1 \;\; \mathcal{F}_X$ thereis some $F_2 \;\; \mathcal{F}_Y$ such that $F_1 \leq F_2$, then it holds that $\overline{F}_X = \sup_{F_1 \; \mathcal{F}_X} F_1 \leq \sup_{F_2 \; \mathcal{F}_Y} F_2 = \overline{F}_Y$.

(2,5) Thesecond (resp. fifth) statement follows from the third (resp., sixth) and Prop osition 4.3. ∎

Nextexampleshows thatthe converse implications inthesecondto sixthstatementsdo nothold in general.

Example 4.20 *Take $\mathcal{F}_X = \{F_{(0.3,0.7),(0,1)}, F_{(0.2,0.8),(0.2,0.3)}\}$, $\mathcal{F}_Y = \{F\}$. They are incomparableunder any ofthe definitions but $\underline{F}_X \leq \underline{F}_Y = F = \overline{F}_Y \leq \overline{F}_X$, from which we deducethat the converse implications in Proposition 4.19 donot hold.*

As we mentioned after Definition 4.1, the differe nce b etween (FSD_2) and (FSD_3) lies on whetherthe set ofdistribution functions $\mathcal{F}_X$ has a "b est case", i.e., a smallest distribution function; similarly, the difference between (FSD_5) and (FSD_6) lies on whether $\mathcal{F}_Y$ has agreatest distribution function. Takingthis intoaccount, we can easilyadapt the conditions of Prop osition 4.6 towards imprecise sto chastic dominance:

Prop osition 4.21 *Let $\mathcal{F}_X$ and $\mathcal{F}_Y$ be twosets of cumulativedistribution functions.*

1. $\underline{F}_X \;\; \mathcal{F}_X \qquad \mathcal{F}_X \;\; _{FSD_2}\; \mathcal{F}_Y \qquad \mathcal{F}_X \;\; _{FSD_3}\; \mathcal{F}_Y$.
2. $\overline{F}_Y \;\; \mathcal{F}_Y \qquad \mathcal{F}_X \;\; _{FSD_5}\; \mathcal{F}_Y \qquad \mathcal{F}_X \;\; _{FSD_6}\; \mathcal{F}_Y$.

Pro of To see the firststatement, use that by Prop osition 4.3 $\mathcal{F}_X \;\; _{FSD_2}\; \mathcal{F}_Y$ implies $\mathcal{F}_X \;\; _{FSD_3}\; \mathcal{F}_Y$. Moreover, $\mathcal{F}_X \;\; _{FSD_3}\; \mathcal{F}_Y$ if and only if for every $F_2 \;\; \mathcal{F}_Y$ there is

$F_1 \in \mathcal{F}_X$ such that $F_1 \le F_2$. In particular, since $\underline{F}_X \le F_1$ for every $F_1 \in \mathcal{F}_X$, it holds that $\underline{F}_X \le F_2$ for every $F_2 \in \mathcal{F}_Y$, and consequently, as $\underline{F}_X \in \mathcal{F}_X$, that $\mathcal{F}_X \succeq_{FSD_2} \mathcal{F}_Y$.

The proof of the second statement is analogous. ∎

When both the lower and upper distributions belong to the corresponding p-box, they can be used to characterise the preferences between them. In that case, the stochastic dominance between two sets of cumulative distribution functions can be characterised by means of the relationships of stochastic dominance between their lower and upper distribution functions.

Corollary 4.22 *Let $\mathcal{F}_X, \mathcal{F}_Y$ be two sets of cumulative distribution functions, and let $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$ be their associated p-boxes. If $\underline{F}_X, \overline{F}_X \in \mathcal{F}_X$ and $\underline{F}_Y, \overline{F}_Y \in \mathcal{F}_Y$, then*

1. $\mathcal{F}_X \succeq_{FSD_1} \mathcal{F}_Y \Leftrightarrow \overline{F}_X \le \underline{F}_Y$.
2. $\mathcal{F}_X \succeq_{FSD_2} \mathcal{F}_Y \Leftrightarrow \mathcal{F}_X \succeq_{FSD_3} \mathcal{F}_Y \Leftrightarrow \underline{F}_X \le \underline{F}_Y$.
3. $\mathcal{F}_X \succeq_{FSD_4} \mathcal{F}_Y \Leftrightarrow \underline{F}_X \le \overline{F}_Y$.
4. $\mathcal{F}_X \succeq_{FSD_5} \mathcal{F}_Y \Leftrightarrow \mathcal{F}_X \succeq_{FSD_6} \mathcal{F}_Y \Leftrightarrow \overline{F}_X \le \overline{F}_Y$.

Proof The first item has already been showed in Proposition 4.19. The equivalences between $(FSD_2) \Leftrightarrow (FSD_3)$ and $(FSD_5) \Leftrightarrow (FSD_6)$ are given by Proposition 4.21. Also, the direct implications of second, third and fourth items are given by Proposition 4.19. Let us prove the converse implications:

- If $\underline{F}_Y \ge \underline{F}_X \in \mathcal{F}_X$, there is some $F_1 \in \mathcal{F}_X$ such that $F_1 \le F_2$ for all $F_2 \in \mathcal{F}_Y$, and as a consequence $\mathcal{F}_X \succeq_{FSD_2} \mathcal{F}_Y$.
- If $\underline{F}_X \le \overline{F}_Y$, then there exist $F_1 \in \mathcal{F}_X$ and $F_2 \in \mathcal{F}_Y$ such that $F_1 \le F_2$, whence $\mathcal{F}_X \succeq_{FSD_4} \mathcal{F}_Y$.
- If $\overline{F}_X \le \overline{F}_Y$, then since $\overline{F}_Y \in \mathcal{F}_Y$, then there is some $F_2 \in \mathcal{F}_Y$ such that $F_1 \le F_2$ for every $F_1 \in \mathcal{F}_X$, because $F_X \le \overline{F}_X$ for any $F_X \in \mathcal{F}_X$. ∎

In Section 2.1.1 we established a characterisation of stochastic dominance in terms of expectations: Theorem 2.10 assures that given two random variables X and Y, $X \succeq_{FSD} Y$ if and only if $E(u(X)) \ge E(u(Y))$ for every increasing function u. When we compare sets of random variables, we must replace these expectations by lower and upper expectations. For any given set of distribution functions $\mathcal{F}$ and any increasing function $u : [0,1] \to \mathbb{R}$, we shall denote $\underline{E}_{\mathcal{F}}(u) := \inf_{F \in \mathcal{F}} E_{P_F}(u)$ and $\overline{E}_{\mathcal{F}}(u) := \sup_{F \in \mathcal{F}} E_{P_F}(u)$.

Theorem 4.23 *Let us consider two sets of cumulative distribution functions $\mathcal{F}_X$ and $\mathcal{F}_Y$, and let $\mathcal{U}$ be the set of all increasing functions $u : [0,1] \to \mathbb{R}$. The following statements hold:*

1. $\mathcal{F}_X \succeq_{FSD_1} \mathcal{F}_Y \Leftrightarrow \underline{E}_{\mathcal{F}_X}(u) \geq \overline{E}_{\mathcal{F}_Y}(u)$ *for every* $u \in \mathcal{U}$.
2. $\mathcal{F}_X \succeq_{FSD_2} \mathcal{F}_Y \Leftrightarrow \overline{E}_{\mathcal{F}_X}(u) \geq \overline{E}_{\mathcal{F}_Y}(u)$ *for every* $u \in \mathcal{U}$.
3. $\mathcal{F}_X \succeq_{FSD_3} \mathcal{F}_Y \Leftrightarrow \overline{E}_{\mathcal{F}_X}(u) \geq \overline{E}_{\mathcal{F}_Y}(u)$ *for every* $u \in \mathcal{U}$.
4. $\mathcal{F}_X \succeq_{FSD_4} \mathcal{F}_Y \Leftrightarrow \overline{E}_{\mathcal{F}_X}(u) \geq \underline{E}_{\mathcal{F}_Y}(u)$ *for every* $u \in \mathcal{U}$.
5. $\mathcal{F}_X \succeq_{FSD_5} \mathcal{F}_Y \Leftrightarrow \underline{E}_{\mathcal{F}_X}(u) \geq \underline{E}_{\mathcal{F}_Y}(u)$ *for every* $u \in \mathcal{U}$.
6. $\mathcal{F}_X \succeq_{FSD_6} \mathcal{F}_Y \Leftrightarrow \underline{E}_{\mathcal{F}_X}(u) \geq \underline{E}_{\mathcal{F}_Y}(u)$ *for every* $u \in \mathcal{U}$.

Proof

1. First of all, $\mathcal{F}_X \succeq_{FSD_1} \mathcal{F}_Y$ if and only if for every $F_1 \in \mathcal{F}_X$ and $F_2 \in \mathcal{F}_Y$ $F_1 \succeq_{FSD} F_2$. This is equivalent to $E_{P_1}(u) \geq E_{P_2}(u)$, for every $u \in \mathcal{U}$, and every $F_1 \in \mathcal{F}_X$ and $F_2 \in \mathcal{F}_Y$, where P_i is the probability associated with F_i, for $i = 1,2$, and this in turn is equivalent to

$$\underline{E}_{\mathcal{F}_X}(u) = \inf\{E_{P_F}(u) \mid F \in \mathcal{F}_X\} \geq \sup\{E_{P_F}(u) \mid F \in \mathcal{F}_Y\} = \overline{E}_{\mathcal{F}_Y}(u)$$

for every $u \in \mathcal{U}$.

3. If $\mathcal{F}_X \succeq_{FSD_3} \mathcal{F}_Y$, then for every $F_2 \in \mathcal{F}_Y$ there is $F_1 \in \mathcal{F}_X$ such that $F_1 \leq F_2$. Equivalently, for every $F_2 \in \mathcal{F}_Y$ there is $F_1 \in \mathcal{F}_X$ such that $E_{P_1}(u) \geq E_{P_2}(u)$ for every $u \in \mathcal{U}$. Then given $u \in \mathcal{U}$ and $F_2 \in \mathcal{F}_Y$,

$$E_{P_2}(u) \leq \sup\{E_{P_F}(u) \mid F \in \mathcal{F}_X\} = \overline{E}_{\mathcal{F}_X}(u),$$

and consequently

$$\overline{E}_{\mathcal{F}_Y}(u) = \sup\{E_{P_F}(u) \mid F \in \mathcal{F}_Y\} \leq \overline{E}_{\mathcal{F}_X}(u).$$

2. The second statement follows from the third one and from Proposition 4.3.

4. Let us assume that $\mathcal{F}_X \succeq_{FSD_4} \mathcal{F}_Y$. Then, by definition there are $F_1 \in \mathcal{F}_X$ and $F_2 \in \mathcal{F}_Y$ such that $F_1 \leq F_2$, or equivalently, $E_{P_1}(u) \geq E_{P_2}(u)$ for every $u \in \mathcal{U}$. We deduce that

$$\begin{aligned}\overline{E}_{\mathcal{F}_X}(u) &= \sup\{E_{P_F}(u) \mid F \in \mathcal{F}_X\} \geq E_{P_1}(u) \\ &\geq E_{P_2}(u) \geq \inf\{E_{P_F}(u) \mid F \in \mathcal{F}_Y\} = \underline{E}_{\mathcal{F}_Y}(u).\end{aligned}$$

6. If $F_X \succeq_{FSD_6} F_Y$, then for every $F_1 \in \mathcal{F}_X$ there is $F_2 \in \mathcal{F}_Y$ such that $F_1 \leq F_2$. Equivalently, for every $F_1 \in \mathcal{F}_X$, $E_{P_1}(u) \geq E_{P_2}(u)$ for some $F_2 \in \mathcal{F}_Y$ and for every $u \in \mathcal{U}$. Thus, for every $F_1 \in \mathcal{F}_X$ and $u \in \mathcal{U}$,

$$E_{P_1}(u) \geq \inf\{E_{P_F}(u) \mid F \in \mathcal{F}_Y\},$$

and consequently

$$\underline{E}_{\mathcal{F}_X}(u) = \inf\{E_{P_F}(u) \mid F \in \mathcal{F}_X\} \geq \inf\{E_{P_F}(u) \mid F \in \mathcal{F}_Y\} = \underline{E}_{\mathcal{F}_Y}(u).$$

5. Finally, the fifth statement follows from the sixth and from Proposition 4.3. ∎

Remark 4.24 *If we consider the extension of stochastic dominance $\succeq_{FSD_{3,6}}$, that is, $F_X \succeq_{FSD_{3,6}} F_Y$ if and only if $F_X \succeq_{FSD_3} F_Y$ and $F_X \succeq_{FSD_6} F_Y$, it holds that:*

$$F_X \succeq_{FSD_{3,6}} F_Y \Leftrightarrow \begin{array}{l} \underline{F}_X \succeq_{FSD} \underline{F}_Y \text{ and } \overline{F}_X \succeq_{FSD} \overline{F}_Y. \\ \overline{E}_{\mathcal{F}_X}(u) \geq \overline{E}_{\mathcal{F}_Y}(u) \text{ and } \underline{E}_{\mathcal{F}_X}(u) \geq \underline{E}_{\mathcal{F}_Y}(u) \quad \forall u \in \mathcal{U}. \end{array} \tag{4.3}$$

With a similar notation, we can consider $\succeq_{FSD_{2,5}}$, and it holds that $F_X \succeq_{FSD_{2,5}} F_Y$ implies $F_X \succeq_{FSD_{3,6}} F_Y$. Then, from the previous results we deduce that $F_X \succeq_{FSD_{2,5}} F_Y$ also implies the results of Equation (4.3).

Taking into account Equation (2.6), the above implications hold in particular when we replace the set $\mathcal{U}$ by the subset $\mathcal{U}^*$ of increasing and bounded functions $u : [0, 1] \to \mathbb{R}$. This will be useful when comparing random sets by means of stochastic dominance in Section 4.2.1.

Remark 4.25 *Theorem 4.23 shows that the extensions of first degree stochastic dominance to sets of alternatives are related to the comparison of the lower and upper expectations they induce. Taking this idea into account, we may introduce alternative definitions by considering a convex combination of these lower and upper expectations, in a similar way to the Hurwicz criterion [96]:*

$$F_X \succeq_{FSD_H} F_Y \Leftrightarrow \lambda\underline{E}_{\mathcal{F}_X}(u) + (1-\lambda)\overline{E}_{\mathcal{F}_X}(u) \geq \lambda\underline{E}_{\mathcal{F}_Y}(u) + (1-\lambda)\overline{E}_{\mathcal{F}_Y}(u),$$

for all $u \in \mathcal{U}$, where $\lambda \in [0,1]$ plays the role of a pessimistic index. It is not difficult to see that

$$F_X \succeq_{FSD_1} F_Y \Rightarrow F_X \succeq_{FSD_{2,5}} F_Y \Rightarrow F_X \succeq_{FSD_{3,6}} F_Y \Rightarrow F_X \succeq_{FSD_H} F_Y$$

and that the converses do not hold.

When the bounds of the p-boxes belong to the sets of distribution functions, the implications on Theorem 4.23 become equivalences.

Corollary 4.26 *Let $\mathcal{F}_X$ and $\mathcal{F}_Y$ be two sets of cumulative distribution functions, and let $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$ be their associated p-boxes. If $\underline{F}_X, \overline{F}_X \in \mathcal{F}_X$ and $\underline{F}_Y, \overline{F}_Y \in \mathcal{F}_Y$, then:*

1. $\mathcal{F}_X \succeq_{FSD_1} \mathcal{F}_Y \Rightarrow \underline{E}_{\mathcal{F}_X}(u) \geq \overline{E}_{\mathcal{F}_Y}(u)$ *for every* $u \in \mathcal{U}$.
2. $\mathcal{F}_X \succeq_{FSD_2} \mathcal{F}_Y \Leftrightarrow \mathcal{F}_X \succeq_{FSD_3} \mathcal{F}_Y \Rightarrow \overline{E}_{\mathcal{F}_X}(u) \geq \overline{E}_{\mathcal{F}_Y}(u)$ *for every* $u \in \mathcal{U}$.
3. $\mathcal{F}_X \succeq_{FSD_4} \mathcal{F}_Y \Rightarrow \overline{E}_{\mathcal{F}_X}(u) \geq \underline{E}_{\mathcal{F}_Y}(u)$ *for every* $u \in \mathcal{U}$.
4. $\mathcal{F}_X \succeq_{FSD_5} \mathcal{F}_Y \Leftrightarrow \mathcal{F}_X \succeq_{FSD_6} \mathcal{F}_Y \Rightarrow \underline{E}_{\mathcal{F}_X}(u) \geq \underline{E}_{\mathcal{F}_Y}(u)$ *for every* $u \in \mathcal{U}$.

Proof The proof is based on the fact that, since $\underline{F}_X, \overline{F}_X \in \mathcal{F}_X$ and $\underline{F}_Y, \overline{F}_Y \in \mathcal{F}_Y$, then:

$$\underline{E}_{\mathcal{F}_X}(u) = E_{\overline{F}_X}(u), \quad \overline{E}_{\mathcal{F}_X}(u) = E_{\underline{F}_X}(u),$$
$$\underline{E}_{\mathcal{F}_Y}(u) = E_{\overline{F}_Y}(u), \quad \overline{E}_{\mathcal{F}_Y}(u) = E_{\underline{F}_Y}(u).$$

Then, applying Corollary 4.22, the implications directly hold. ∎

It is also possible to consider the n-th degree stochastic dominance, for $n \geq 2$ as the binary relation in Definition 4.1. In that case, we shall denote by $\succeq_{nSD_i}$ or by (nSD_i) its extensions. With this relation, we can also state similar results to the ones established for first degree stochastic dominance. For instance, the following statements hold for imprecise n-th degree stochastic dominance:

- $\mathcal{F}_X \succeq_{nSD_2} \mathcal{F}_Y \Leftrightarrow \mathcal{F}_X \succeq_{nSD_3} \mathcal{F}_Y$ (the proof is analogous to that of Proposition 4.15).
- $\mathcal{F}_X \succeq_{nSD_5} \mathcal{F}_Y \Leftrightarrow \mathcal{F}_X \succeq_{nSD_6} \mathcal{F}_Y$ (the proof is analogous to that of Proposition 4.15).
- $\mathcal{F}_X \succeq_{nSD_i} \mathcal{F}_Y \Rightarrow \mathcal{F}_X \succeq_{mSD_i} \mathcal{F}_Y$ for any $n<m$ (see Equation (2.4)).

In addition, the connection of the comparison of sets of cumulative distribution functions with the associated p-boxes (Proposition 4.19) or with the associated lower and upper expectations (Theorem 4.23) can also be stated for the imprecise n-th degree stochastic dominance as follows:

Proposition 4.27 *Let $\mathcal{F}_X$ and $\mathcal{F}_Y$ be two sets of cumulative distribution functions, and denote by $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$ the associated p-boxes. Denote by $\mathcal{U}_n$ the set of bounded and increasing functions $u : \mathbb{R} \to \mathbb{R}$ that are n-monotone. Then it holds that:*

- $\mathcal{F}_X \succeq_{nSD_1} \mathcal{F}_Y$ *holds if and only if* $\overline{F}_X \succeq_{nSD_1} \underline{F}_Y$, *and this is equivalent to*

$$\underline{E}_{\mathcal{F}_X}(u) \geq \overline{E}_{\mathcal{F}_Y}(u)$$

for every $u \in \mathcal{U}_n$.

- $\mathcal{F}_X \succeq_{nSD_2} \mathcal{F}_Y$ *implies:*

$$\underline{F}_X \succeq_{nSD_2} \underline{F}_Y \text{ and } \overline{E}_{\mathcal{F}_X}(u) \geq \overline{E}_{\mathcal{F}_Y}(u) \text{ for every } u \in \mathcal{U}_n.$$

- $\mathcal{F}_X \succeq_{nSD_3} \mathcal{F}_Y$ *implies:*

$$\underline{F}_X \succeq_{nSD_3} \underline{F}_Y \text{ and } \overline{E}_{\mathcal{F}_X}(u) \geq \overline{E}_{\mathcal{F}_Y}(u) \text{ for every } u \in \mathcal{U}_n.$$

- $\mathcal{F}_X \succeq_{nSD_4} \mathcal{F}_Y$ *implies:*

$$\underline{F}_X \succeq_{FSD_4} \overline{F}_Y \text{ and } \overline{E}_{\mathcal{F}_X}(u) \geq \underline{E}_{\mathcal{F}_Y}(u) \text{ for every } u \in \mathcal{U}_n.$$

- $\mathcal{F}_X \succeq_{nSD_5} \mathcal{F}_Y$ *implies:*

$$\overline{F}_X \succeq_{nSD_5} \overline{F}_Y \text{ and } \underline{E}_{\mathcal{F}_X}(u) \geq \underline{E}_{\mathcal{F}_Y}(u) \text{ for every } u \in \mathcal{U}_n.$$

- $\mathcal{F}_X \succeq_{nSD_6} \mathcal{F}_Y$ *implies:*

$$\overline{F}_X \succeq_{nSD_6} \overline{F}_Y \text{ and } \underline{E}_{\mathcal{F}_X}(u) \geq \underline{E}_{\mathcal{F}_Y}(u) \text{ for every } u \in \mathcal{U}_n.$$

Furthermore, the converse implications hold when $\underline{F}_X, \overline{F}_X \in \mathcal{F}_X$ *and* $\underline{F}_Y, \overline{F}_Y \in \mathcal{F}_Y$.

We omit the proof because it is analogous to the one of Proposition 4.19, Theorem 4.23 and Corollaries 4.22 and 4.26.

In the remainder of the subsection we shall investigate several properties of imprecise stochastic dominance. However, from now on we shall focus on the first degree stochastic dominance for two main reasons: on the one hand, it is the most common stochastic dominance in the literature and, on the other hand, as we have just seen, the results for first degree can be easily extended for n-th degree stochastic dominance.

Connection with previous approaches

A first approach to the extension of the stochastic dominance towards an imprecise framework was made by Denoeux in [61].

He considered two random variables U and V such that $P(U \leq V) = 1$. They can be equivalently represented as a random interval $[U, V]$, which in turn induces a belief and a plausibility function, as we saw in Definition 2.43:

$$bel(A) = P([U, V] \subseteq A) \text{ and } pl(A) = P([U, V] \cap A \neq \emptyset)$$

for every element A in the Borel sigma-algebra $\beta_{\mathbb{R}}$. Thus, for every $x \in \mathbb{R}$:

$$bel((-\infty, x]) = F_V(x) \text{ and } pl((-\infty, x]) = F_U(x).$$

The asso ciated set of probability measures P compatible with bel and pl is give n by:

$$P = \{P \text{ probability } : bel(A) \leq P(A) \leq pl(A) \text{ for every } A \quad \beta_R\}.$$

Deno eux considered two random closed intervals $[U, V]$ and $[U, V]$. One possible way of comparing them is to compare their asso ciated sets of probabilities:

$$P = \{P \text{ probability } : bel(A) \leq P(A) \leq pl(A) \text{ for every } A \quad \beta_R\}.$$
$$P = \{P \text{ probability } : bel(A) \leq P(A) \leq pl(A) \text{ for every } A \quad \beta_R\}.$$

Based onthe usual ordering b etwee n realintervals (see [78]), Denoeux prop osed the following notions:

- $P \quad P \quad pl((x, \infty)) \leq bel((x, \infty))$ for every $x \quad \mathbb{R}$.
- $P \quad P \quad pl((x, \infty)) \leq pl((x, \infty))$ for every $x \quad \mathbb{R}$.
- $P \quad P \quad bel((x, \infty)) \leq bel((x, \infty))$ for every $x \quad \mathbb{R}$.
- $P \quad P \quad bel((x, \infty)) \leq pl((x, \infty))$ for every $x \quad \mathbb{R}$.

It turns out that the above notions can be characterised in terms of the sto chastic dominance b etween the lower and upp er limits of the random intervals:

Prop osition 4.28 ([61]) *Let (U, V) and (U, V) be two pairs of random variables satisfying $P(U \leq V) = P(U \leq V) = 1$, and let P and P their associated sets of probability measures. The fol lowing equivalences hold:*

- $P \quad P \quad U \ {}_{\text{FSD}}\ V$.
- $P \quad P \quad U \ {}_{\text{FSD}}\ U$.
- $P \quad P \quad V \ {}_{\text{FSD}}\ V$.
- $P \quad P \quad V \ {}_{\text{FSD}}\ U$.

Note that the ab ove definitions can be represented in an equivalent way by means of p-b oxes:if we conside r the set of distribution functions induced by P, we obtain

$$\{F : F_V \leq F \leq F_U\},$$

i.e., the p-b ox determined by F_V and F_U. Similarly, the set P induces the p-b ox (F_V, F_U), and Deno eux's definitions are equivalent to comparing the lower and upp er distribution functions of these p-b oxes, as we can see from Prop osition 4.28. Note moreover thatthe same result holds if we considerfinitelyadditive probability measures instead of σ-additive ones, b ecause b oth of them determine the same p-b ox and the lower and upp er distribution functions are included in both cases.

There is a clear connection b etween the scenario prop osed by Deno eux and our proposal. Let $[U, V]$ and $[U, V]$ b e two random closed intervals, whose asso ciated belief and plausibility functions determine the setsofprobability measures P, P andthe setsofcumulative distribution functions F and F . Applying Prop osition 4.28 and Corollary 4.22, we obtain the following equivalences:

- $F \ _{FSD_1} \ F \quad F_U(t) \leq F_V(t)$ for every $t \ \mathbb{R} \ P \quad P$.
- $F \ _{FSD_2} \ F \quad F \ _{FSD_3} \ F \quad F_V(t) \leq F_V(t)$ for every $t \ \mathbb{R} \ P \quad P$.
- $F \ _{FSD_4} \ F \quad F_V(t) \leq F_U(t)$ for every $t \ \mathbb{R} \ P \quad P$.
- $F \ _{FSD_5} \ F \quad F \ _{FSD_6} \ F \quad F_U(t) \leq F_U(t)$ for every $t \ \mathbb{R} \ P \quad P$.

Hence, condition gives rise to (FSD_2) (when P has a smallest distribution function) and (FSD_3) (when it do es not have it); similarly, condition pro duces (FSD_5) (if P has a greatest distributionfunction) and (FSD_6) (otherwise).

This also shows that our prop osal is more general in the sense that it can be applied to arbitrary sets of probability measures, and not only those asso ciated with a random closed interval. Ontheotherhand, ourworkismorerestrictiveinthesensethatweare assuming that our referential space is [0,1], instead of the real line. Aswe mentioned at the beginning of the section, our results are imm ediately extendable to distribution functions taking values in any closedinterval $[a, b]$ where $a<b$ are real numb ers. The restriction to b ounded intervals is made so that the lower envelop e of a set of cumulative distribution functions is a finitely additive distribution function, which may not be the case if we consider the whole real line as our referential space. Onesolution tothis problem is to add to our space a smallest and a greatest value $0_\Omega, 1_\Omega$, so that we always have $F(0_\Omega) = 0$ and $F(1_\Omega) = 1$.

Increasing imprecision

Next we study the behaviour of the differentnotions of sto chastic dominance for sets of distributions when we use them to compare two sets of distribution functions, one of which is more imprecis e than the other. This may be useful in some situations: for instance, p-b oxes can be seen as *confidence bands* [87,374], which mo delour imprecise information ab out a distribution function taking into account a given sample and a fixed confidence level. Thenifweapplytwodifferentconfidence levelsto thesamedata, we obtain two con fidence bands, one included in the other, and we may study which of the two is preferred according to the different criteria we have prop osed. In thissense, we may also study our preferences b etween a set of portfolios, that we representbymeans of a set of distribution functions, and a greater set, where we include more distribution functions, but where also the asso ciated risk may increase.

Wearegoing toconsider two different situations: thefirstone iswhenourinformation is given by a set of distribution fu nctions. Hence, weconsidertwosets $F_X \quad F_Y$ and investigate our preferences b etween them:

Prop osition 4.29 *Let us consider two set s of cumulative distribution functions F_X and F_Y such that $F_X \quad F_Y$. It holdsthat:*

1. *If F_X hasonly onedistributionfunction, then all thepossibilities arevalid for (FSD_1). Otherwise, if F_X is formedby more than one distributionfunction, F_X and F_Y are incomparable withrespect to (FSD_1).*
2. *With respectto $(FSD_2), \ldots, (FSD_6)$, thepossible scenariosare summarisedin the fol lowing table:*

	FSD_2	FSD_3	FSD_4	FSD_5	FSD_6
$F_X \quad_{FSD_i} F_Y$				•	•
$F_Y \quad_{FSD_i} F_X$	•	•			
$F_X \equiv_{FSD_i} F_Y$	•	•	•	•	•
F_X, F_Y incomparable	•			•	

Pro of Let us prove that the p ossibilities ruled out in the statement of the prop osition cannot happ en:

1. Onthe onehand, if F_X has morethanone cumulative distribution function, we deduce that F_X is in comparable with i tself with resp ect to (FSD_1), andasa consequence it is also incomparable with resp ect to the greater set F_Y.
2. Since $F_X \quad F_Y$, for any $F_1 \quad F_X$ there exists $F_2 \quad F_Y$ such that $F_1 = F_2$. Hence, we always have $F_Y \quad_{FSD_3} F_X$ and $F_X \quad_{FSD_6} F_Y$. Thus, weobtainthat $F_X \quad_{FSD_3} F_Y$, $F_Y \quad_{FSD_6} F_X$, and b oth sets cannot be incomparable with resp ect to (FSD_3) and (FSD_6). Moreover, using Prop osition 4.3 $F_X \quad_{FSD_2} F_Y$ and $F_Y \quad_{FSD_5} F_X$ are not possible. Thisalsoshows that $F_X \equiv_{FSD_4} F_Y$, because any $F \quad F_X \quad F_Y$ is equivalent to itself. ■

Next example shows that all the other scenarios are indeed possible.

Example 4.30 • *Let ussee that $F_X \quad_{FSD_i} F_Y$ is possible for $i = 1, 5, 6$. For this aim, take $F_X = \{F\}$ and $F_Y = \{F, F_{1,0}\}$. Then, it holds that $F_X \quad_{FSD_i} F_Y$ for $i = 1, 5, 6$ and $F_X \equiv_{FSD_i} F_Y$ for $i = 2, 3$.*

- *Let us check that* $F_Y \ {}_{FSD_i} F_X$, *is possible for* $i = 1, 2, 3$. *Consider* $F_X = \{F\}$ *and* $F_Y = \{F, F_{1,1}\}$. *Then, it holds that* $F_Y \ {}_{FSD_i} F_X$ *for* $i = 1, 2, 3$ *and* $F_X \equiv_{FSD_i} F_Y$ *for* $i = 5, 6$.
- *Now, let us see that* $F_X \equiv_{FSD_i} F_Y$, *is possible for* $i = 1, \ldots, 6$. *For this aim, take* $F_X = F_Y = \{F\}$. *Then,* $F_X \equiv_{FSD_1} F_Y$ *and by Proposition 4.3,* $F_X \equiv_{FSD_i} F_Y$ *for any* $i = 2, \ldots, 6$.
- *To see that incomparability is possible for* $i = 1, 2, 5$, *let* $F_X = F_Y = \{F, F_{1,0.5}\}$. *Then* F_X *and* F_Y *are* (FSD_i) *incomparable for* $i = 1, 2, 5$, *since* F *and* $F_{1,0.5}$ *are incomparable.*

Remark 4.31 *A particular case of the above result would be when we compare a set of distribution functions* F_X *with itself, i.e., when* $F_Y = F_X$. *In that case,* $F_X \equiv_{FSD_i} F_X$ *for* $i = 3, 4, 6$, *as we have seen in Proposition 4.9. With respect to* (FSD_1), (FSD_2) *and* (FSD_5), *we may have either incomparability or indifference: to see that we may have incomparability, consider* $F_X = F_Y = \{F, F_{1,0.5}\}$; *for indifference take* $F_X = F_Y = \{F\}$.

The second scenario corresponds to the case where our information about the set of distribution functions is given by means of a p-box. A more imprecise p-box corresponds to the case where either the lower distribution function is smaller, the upper distribution function is greater, or both. We begin by considering the latter case.

Proposition 4.32 *Let us consider two sets of cumulative distribution functions* F_X *and* F_Y, *and let* $(\underline{F}_X, \overline{F}_X)$ *and* $(\underline{F}_Y, \overline{F}_Y)$ *denote their associated p-boxes. Assume that* $\underline{F}_Y < \underline{F}_X < \overline{F}_X < \overline{F}_Y$. *Then the possible scenarios of stochastic dominance are summarised in the following table:*

	FSD_1	FSD_2	FSD_3	FSD_4	FSD_5	FSD_6
$F_X \ {}_{FSD_i} F_Y$				•	•	•
$F_Y \ {}_{FSD_i} F_X$		•	•	•		
$F_X \equiv_{FSD_i} F_Y$				•		
F_X, F_Y *incomparable*	•	•	•	•	•	•

Proof Using Proposition 4.3, we know that $F_X \ {}_{FSD_1} F_Y$ if and only if $\overline{F}_X \leq \underline{F}_Y$, which is incompatible with the assumptions. Similarly, we can see that $F_Y \ {}_{FSD_1} F_X$ and as a consequence they are incomparable.

On the other hand, if $F_X \ {}_{FSD_i} F_Y$, for $i = 2, 3$, using Proposition 4.19 it holds that $\underline{F}_X \leq \underline{F}_Y$, a contradiction with the hypothesis.

Similarly, if $F_Y \ {}_{FSD_i} F_X$, for $i = 5, 6$, we deduce from Proposition 4.19 that $\overline{F}_Y \leq \overline{F}_X$, again a contradiction. ∎

Next example shows that the scenarios included in the table are p oss ible.

Example 4.33 • *Letusseethat for* (FSD_i), $i = 2,\ldots,6$, $\mathcal{F}_X$ *and* $\mathcal{F}_Y$ *can be incomparable. For thisaim weconsider* $\mathcal{F}_X = \{F, F\}$*,where* F $=\max\{F, F_{1,0.7}\}$*, and* $\mathcal{F}_Y = \{F_{1,0.5}, F_{\{(0.5,0.5),(0,1)\}}\}$*. It is easy to check that bot h sets of cumulative distribution functionsare incomparable, since every distribution functionon* $\mathcal{F}_X$ *is incomparable with every distribution function on* $\mathcal{F}_Y$*.*

- *Let usnow consider*

$$\mathcal{F}_X = \{F, F\} \text{ and } \mathcal{F}_Y = \{F_{(0.5,0.5),(0,0.5)}, F_{(0.5,0.5),(0.5,1)}\}.$$

Then $\mathcal{F}_Y \quad_{FSD_i} \mathcal{F}_X$ *for* $i = 2,3$ *and* $\mathcal{F}_X \quad_{FSD_i} \mathcal{F}_Y$ *for* $i = 5,6$ *. As a consequence, both sets are indifferent with respect to Definition* (FSD_4)*.*

- *Final ly, it only remains to see that we may have strict preference under Definition* (FSD_4)*. On theone hand, ifwe considerthe sets*

$$\mathcal{F}_X = \{F, F\} \text{ and } \mathcal{F}_Y = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}, F_{(0.5,0.5),(0,0.5)}\},$$

it holds that $\mathcal{F}_X \quad_{FSD_4} \mathcal{F}_Y$*. In theother hand, ifwe consider*

$$\mathcal{F}_Y = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}, F_{(0.5,0.5),(0.5,1)}\},$$

we obtain that $\mathcal{F}_Y \quad_{FSD_4} \mathcal{F}_X$*.*

Although the inclusion $\mathcal{F}_X \quad \mathcal{F}_Y$ implies that $\underline{F}_Y \leq \underline{F}_X \leq \overline{F}_X \leq \overline{F}_Y$, wemay have $\underline{F}_Y < \underline{F}_X < \overline{F}_X < \overline{F}_Y$ even if $\mathcal{F}_X$ and $\mathcal{F}_Y$ aredisjoint, forinstancewhentheselowerand upp er distribution functions are σ-additive and we take the sets $\mathcal{F}_X = \{\underline{F}_X, \overline{F}_X\}$ and $\mathcal{F}_Y = \{\underline{F}_Y, \overline{F}_Y\}$. For this reason in Prop osition 4.29 we cannot have $\mathcal{F}_X \quad_{FSD_4} \mathcal{F}_Y$ nor $\mathcal{F}_Y \quad_{FSD_4} \mathcal{F}_X$ and under the conditions of Prop osition 4.32 we can.

Prop osition 4.34 *Undertheaboveconditions, ifinaddition* $\underline{F}_X, \overline{F}_X$ *belong to* $\mathcal{F}_X$ *and* $\underline{F}_Y, \overline{F}_Y$ *belong to* $\mathcal{F}_Y$*, the possible scenarios are:*

	FSD_1	FSD_2	FSD_3	FSD_4	FSD_5	FSD_6
$F_1 \quad_{FSD_i} F_2$					•	•
$F_2 \quad_{FSD_i} F_1$		•	•			
$F_1 \equiv_{FSD_i} F_2$				•		
F_1, F_2 *incomparable*	•					

Pro of

- It isobvious that $\mathcal{F}_X$ and $\mathcal{F}_Y$ are incomparable with resp ect to Definition (FSD_1).

- It holds that $\underline{F}_Y < \underline{F}_X \leq F_1$ for any $F_1 \in \mathcal{F}_X$, and then $\mathcal{F}_Y \succeq_{FSD_2} \mathcal{F}_X$. Moreover, using Corollary 4.22 (FSD_2) and (FSD_3) are equivalent, and consequently $\mathcal{F}_Y \succeq_{FSD_3} \mathcal{F}_X$.
- We know that $\underline{F}_Y < \underline{F}_X$, then $\mathcal{F}_Y \succeq_{FSD_4} \mathcal{F}_X$, and moreover $\overline{F}_X < \overline{F}_Y$, and then $\mathcal{F}_X \succeq_{FSD_4} \mathcal{F}_Y$. Using both inequalities we obtain that $\mathcal{F}_X \equiv_{FSD_4} \mathcal{F}_Y$.
- It holds that $F_1 \leq \overline{F}_X < \overline{F}_Y$ for any $F_1 \in \mathcal{F}_X$, and then $\mathcal{F}_X \succeq_{FSD_5} \mathcal{F}_Y$. Furthermore, using Corollary 4.22, (FSD_5) and (FSD_6) are equivalent, and consequently $\mathcal{F}_X \succeq_{FSD_6} \mathcal{F}_Y$. ∎

In particular, the above result is applicable when $\mathcal{F}_X = (\underline{F}_X, \overline{F}_X)$ and $\mathcal{F}_Y = (\underline{F}_Y, \overline{F}_Y)$, with $\underline{F}_X, \overline{F}_X \in \mathcal{F}_X$ and $\underline{F}_Y, \overline{F}_Y \in \mathcal{F}_Y$.

To conclude this part, we consider the case where only one of the bounds becomes more imprecise in the second p-box.

Proposition 4.35 *Let us consider two sets of cumulative distribution functions $\mathcal{F}_X$ and $\mathcal{F}_Y$, and let $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$ denote their associated p-boxes.*

a) Let us assume that $\underline{F}_Y < \underline{F}_X < \overline{F}_X = \overline{F}_Y$. Then the possible scenarios are:

	FSD_1	FSD_2	FSD_3	FSD_4	FSD_5	FSD_6
$\mathcal{F}_X \succ_{FSD_i} \mathcal{F}_Y$				•	•	•
$\mathcal{F}_Y \succ_{FSD_i} \mathcal{F}_X$		•	•	•	•	•
$\mathcal{F}_X \equiv_{FSD_i} \mathcal{F}_Y$				•	•	•
$\mathcal{F}_X, \mathcal{F}_Y$ *incomparable*	•	•	•	•	•	•

b) Let us assume that $\underline{F}_Y = \underline{F}_X < \overline{F}_X < \overline{F}_Y$. Then the possible situations are:

	FSD_1	FSD_2	FSD_3	FSD_4	FSD_5	FSD_6
$\mathcal{F}_X \succ_{FSD_i} \mathcal{F}_Y$		•	•	•	•	•
$\mathcal{F}_Y \succ_{FSD_i} \mathcal{F}_X$		•	•	•		
$\mathcal{F}_X \equiv_{FSD_i} \mathcal{F}_Y$		•	•	•		
$\mathcal{F}_X, \mathcal{F}_Y$ *incomparable*	•	•	•	•	•	•

Proof

a) Let us first show that incomparability is the only situation possible according to Definition (FSD_1). As proven in Proposition 4.19, $\mathcal{F}_X \succeq_{FSD_1} \mathcal{F}_Y$ if and only if $\overline{F}_X \leq \underline{F}_Y$. But this inequality is not compatible with the hypothesis. For the same reason, the converse inequality, $\overline{F}_Y \leq \underline{F}_X$ is not possible either.

With respect to (FSD_2), (FSD_3), note that if $\underline{F}_Y < \underline{F}_X$,

$$\exists x_0 \in [0,1] \text{ such that } \underline{F}_Y(x_0) = \inf_{F_2 \in \mathcal{F}_Y} F_2(x_0) < \underline{F}_X(x_0)$$

whence there exists $F_2 \in \mathcal{F}_Y$ such that $F_2(x_0) < \underline{F}_X(x_0) \le F_1(x_0)$ for all $F_1 \in \mathcal{F}_X$. Thus, $F_1 \not\le F_2$ for any $F_1 \in \mathcal{F}_X$ and $\mathcal{F}_X \not\succeq_{FSD_3} \mathcal{F}_Y$. Applying Proposition 4.19, $\mathcal{F}_X \not\succeq_{FSD_2} \mathcal{F}_Y$.

b) The proof concerning Definition (FSD_1) is analogous to the one in a).
Concerning (FSD_5), (FSD_6), note that since $\overline{F}_X < \overline{F}_Y$,

$$\exists x_0 \in [0,1] \text{ such that } \overline{F}_Y(x_0) = \sup_{F_2 \in \mathcal{F}_Y} F_2(x_0) > \overline{F}_X(x_0),$$

whence there is $F_2 \in \mathcal{F}_Y$ such that $F_2(x_0) > \overline{F}_X(x_0) \ge F_1(x_0)$ for all $F_1 \in \mathcal{F}_X$, then $F_1 \not\ge F_2$ for any $F_1 \in \mathcal{F}_X$ and $\mathcal{F}_Y \not\succeq_{FSD_6} \mathcal{F}_X$. It also follows from Proposition 4.19 that $\mathcal{F}_Y \not\succeq_{FSD_5} \mathcal{F}_X$. ∎

Next we give examples showing that when the lower distribution function is smaller in the second p-box and the upper distribution functions coincide, all the possibilities not ruled out in the first table of the previous proposition can arise. Similar examples can be constructed for the case where $\underline{F}_X = \underline{F}_Y$ and $\overline{F}_X < \overline{F}_Y$.

Example 4.36 • *We begin by showing that $\mathcal{F}_X$ and $\mathcal{F}_Y$ can be incomparable under any definition (FSD_i) for $i = 2, \ldots, 6$. Let us consider the sets:*

$$\mathcal{F}_X = \{F_{(0.5-\frac{1}{n},0.5,\frac{1}{n}),(0,0.5,1)} \mid n \ge 3\} \text{ and } \mathcal{F}_Y = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\}.$$

For all $F_1 \in \mathcal{F}_X$ and $F_2 \in \mathcal{F}_Y$ it holds that $F_2 \not\succeq_{FSD} F_1$ and $F_1 \not\succeq_{FSD} F_2$. Then, $\mathcal{F}_X$ and $\mathcal{F}_Y$ are incomparable according to (FSD_4), and therefore also according to (FSD_i) for $i = 2, 3, 5, 6$.

- *To see that $\mathcal{F}_X, \mathcal{F}_Y$ can be indifferent according to (FSD_4), (FSD_5) or (FSD_6), take:*

$$\mathcal{F}_X = \{F_{1,0.5}, F_{(0.5,0.5),(0,0.5)}\} \text{ and } \mathcal{F}_Y = \{F_{(0.5,0.5),(0,0.5)}, F_{1,1}\}.$$

Since $\overline{F}_X = \overline{F}_Y = F_{(0.5,0.5),(0,0.5)}$ belongs to both sets, they verify that $\mathcal{F}_X \succeq_{FSD_5} \mathcal{F}_Y$ and also $\mathcal{F}_Y \succeq_{FSD_5} \mathcal{F}_X$. Therefore, $\mathcal{F}_X \equiv_{FSD_5} \mathcal{F}_Y$. As a consequence, they are also indifferent according to (FSD_6) and (FSD_4).

- *Next we show that it is also possible that $\mathcal{F}_X \succ_{FSD_i} \mathcal{F}_Y$ for $i = 5, 6$. Let us consider*

$$\mathcal{F}_X = \{F_{(1-\frac{1}{n},\frac{1}{n}),(0,1)} : n \ge 3\} \text{ and } \mathcal{F}_Y = \{F_{1,0}, F_{1,1}\}.$$

They verify that $\mathcal{F}_X \succeq_{FSD_5} \mathcal{F}_Y$ since $F_{(1-\frac{1}{n},\frac{1}{n}),(0,1)} \succeq_{FSD} F_{1,0}$ for all n; but $\mathcal{F}_Y \not\succeq_{FSD_5} \mathcal{F}_X$ since there is not $F \in \mathcal{F}_X$ such that $F_{1,0} \succeq_{FSD} F$. We conclude that $\mathcal{F}_X \succ_{FSD_5} \mathcal{F}_Y$, and applying Proposition 4.15 also $\mathcal{F}_X \succ_{FSD_6} \mathcal{F}_Y$.

- *To see that we may also have* $F_Y \quad {}_{FSD_i} F_X$ *for* $i = 5,6$*, take:*

$$F_X = \{F_{1,0}, F_{(0.75,0.25),(0,1)}\} \text{ and } F_Y = \{F_{(1-\frac{1}{n},\frac{1}{n}),(0,1)} : n \geq 3\}.$$

Then $F_Y \quad {}_{FSD_5} F_X$ *because* $F_{(1-\frac{1}{n},\frac{1}{n}),(0,1)} \quad {}_{FSD} F_{1,0}$ *for every n, but they are not indifferent with respect to* (FSD_5)*. Hence,* $F_Y \quad {}_{FSD_5} F_X$ *and applying Proposition 4.15 also* $F_Y \quad {}_{FSD_6} F_X$*.*

- *Let us give next an example where* $F_X \quad {}_{FSD_4} F_Y$*. Consider*

$$F_X = \{F_{(0.6,0.4),(0.5,1)}, F_{(0.5-\frac{1}{n},0.5,\frac{1}{n}),(0,0.5,1)} : n \geq 3\} \text{ and}$$
$$F_Y = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\}.$$

Then, $F_X \quad {}_{FSD_4} F_Y$ *since* $F_{(0.6,0.4),(0.5,1)} \quad {}_{FSD} F_{1,0.5}$ *but* $F_Y \quad {}_{FSD_4} F_X$ *since*

$$F_{1,0.5}(0.5) > F_{(0.5-\frac{1}{n},0.5,\frac{1}{n}),(0,0.5,1)}(0.5) \text{ for all } n \geq 3$$

and $F_{1,0.5}(0.5) > F_{(0.6,0.4),(0.5,1)}(0.5)$*. Also*

$$F_{(0.5,0.5),(0,1)}(0) > F_{(0.5-\frac{1}{n},0.5,\frac{1}{n}),(0,0.5,1)}(0) \text{ for all } n \geq 3$$

and $F_{(0.5,0.5),(0,1)}(0) > F_{(0.6,0.4),(0.5,1)}(0)$.

- *We conclude by showing that it may also happen that* $F_Y \quad {}_{FSD_i} F_X$ *for* $i = 2,3,4$*. Let us consider*

$$F_X = \{F_{(0.5-\frac{1}{n},0.5,\frac{1}{n}),(0,0.5,1)} : n \geq 3\} \text{ and}$$
$$F_Y = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}, F_{(0.5,0.5),(0.5,1)}\}.$$

It holds that

$$F_{(0.5,0.5),(0.5,1)} \quad {}_{FSD} F_{(0.5-\frac{1}{n},0.5,\frac{1}{n}),(0,0.5,1)} \text{ for all } n \geq 3,$$

whence $F_Y \quad {}_{FSD_i} F_X$ *for* $i = 2,3,4$*. On the other hand,*

$$F_{(0.5-\frac{1}{n},0.5,\frac{1}{n}),(0,0.5,1)}(0) > F_{(0.5,0.5),(0.5,1)}(0)$$

and

$$F_{(0.5-\frac{1}{n},0.5,\frac{1}{n}),(0,0.5,1)}(0.5) > F_{(0.5,0.5),(0.5,1)}(0.5),$$

whence $F_X \quad {}_{FSD_i} F_Y$ *for* $i = 2,3,4$*.*

Sets of distribution functions associated with the same p-box

Next we investigate the relationships between the preferences on two sets of distributions functions associated with the same p-box. We consider the case of non-trivial p-boxes (that is, those where the lower and the upper distribution functions are different), since otherwise we obviously obtain indifference.

Proposition 4.37 *Let us consider two sets of cumulative distribution functions* $\mathcal{F}_X$ *and* $\mathcal{F}_Y$ *such that* $\underline{F}_X = \underline{F}_Y$, $\overline{F}_X = \overline{F}_Y$ *and* $\underline{F}_X < \overline{F}_X$. *Then:*

1. $\mathcal{F}_X$ *and* $\mathcal{F}_Y$ *are incomparable with respect to* FSD_1.
2. *With respect to* (FSD_i), $i = 2, \ldots, 6$, *we may have incomparability, strict stochastic dominance or indifference between* $\mathcal{F}_X$ *and* $\mathcal{F}_Y$.

Proof By Proposition 4.19, $\mathcal{F}_X \ {}_{FSD_1}\ \mathcal{F}_Y$ if and only if $\overline{F}_X \leq \underline{F}_Y$, which in this case holds if and only if $\overline{F}_X = \underline{F}_X$, a contradiction with our hypotheses. ∎

With respect to conditions $(FSD_2), \ldots, (FSD_6)$, it is easy to find examples of indifference by taking $\mathcal{F}_X = \mathcal{F}_Y$ including the lower and upper distribution functions. Next example shows that we may also have strict dominance or incomparability.

Example 4.38 *In these examples we are going to show that, given two sets of cumulative distribution functions* $\mathcal{F}_X$ *and* $\mathcal{F}_Y$ *associated with the same p-box, then there can be strict dominance or incomparability (that they may also be indifferent has already been showed in Proposition 4.37).*

- *Let us consider*

  $$\mathcal{F}_X = \{F_{(0.5,0.5),(0,0.5)}, F_{(0.5,0.5),(0.5,1)}\} \text{ and } \mathcal{F}_Y = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\}.$$

  Then, it holds that $\mathcal{F}_X \ {}_{FSD_i}\ \mathcal{F}_Y$ *for* $i = 2,3$ *and* $\mathcal{F}_Y \ {}_{FSD_i}\ \mathcal{F}_X$ *for* $i = 5,6$. *By reversing the roles of* $\mathcal{F}_X$ *and* $\mathcal{F}_Y$, *we obtain an example of* $\mathcal{F}_X$ *and* $\mathcal{F}_Y$ *inducing the same p-box and with* $\mathcal{F}_X \ {}_{FSD_i}\ \mathcal{F}_Y$ *for* $i = 5,6$ *and* $\mathcal{F}_Y \ {}_{FSD_i}\ \mathcal{F}_X$ *for* $i = 2,3$.

- *To see the incomparability, take*

  $$\mathcal{F}_X = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\} \text{ and}$$
  $$\mathcal{F}_Y = \{F_{(\frac{1}{n},0.5,0.5-\frac{1}{n}),(0,0.5,1)}, F_{(0.5-\frac{1}{n},0.5,\frac{1}{n}),(0,0.5,1)} : n \geq 3\}.$$

  It is easy to check that both sets are incomparable with respect to (FSD_4), *and then they are also incomparable with respect to* (FSD_i) *for* $i = 1, \ldots, 6$.

- *Finally, if we consider* $\mathcal{F}_X = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\}$ *and*

  $$\mathcal{F}_Y = \{F_{(\frac{1}{n},0.5,0.5-\frac{1}{n}),(0,0.5,1)}, F_{(0.5-\frac{1}{n},0.5,\frac{1}{n}),(0,0.5,1)} : n \geq 3, F_{(0.5,0.5),(0.5,1)}\}.$$

  We have that $F_{(0.5,0.5),(0.5,1)} \ {}_{FSD}\ F_{1,0.5}$, *while none of the distribution functions in* $\mathcal{F}_X$ *is dominated by a distribution function in* $\mathcal{F}_Y$. *Thus,* $\mathcal{F}_Y \ {}_{FSD_4}\ \mathcal{F}_X$. *Again, reversing the roles of* $\mathcal{F}_X$ *and* $\mathcal{F}_Y$ *we see that we can also have* $\mathcal{F}_Y \ {}_{FSD_4}\ \mathcal{F}_X$.

When the lower and upper distribution functions belong to our set of distributions, we deduce the following result.

Corollary 4.39 *Let us consider two sets of cumulative distribution functions $\mathcal{F}_X$ and $\mathcal{F}_Y$ such that $\underline{F}_X = \underline{F}_Y$, $\overline{F}_X = \overline{F}_Y$, $\underline{F}_X < \overline{F}_X$ and $\underline{F}_X, \overline{F}_X \in \mathcal{F}_X \cap \mathcal{F}_Y$. Then $\mathcal{F}_X \equiv_{FSD_i} \mathcal{F}_Y$ for $i = 2, \ldots, 6$, and they are incomparable with respect to (FSD_1).*

Proof The result follows immediately from Proposition 4.37 and Corollary 4.22. ∎

Next we investigate the case where we compare these two sets of distribution functions with a third one, and determine if they produce the same preferences:

Proposition 4.40 *Let us consider $\mathcal{F}_X$, $\mathcal{F}'_X$ and $\mathcal{F}_Y$ three sets of cumulative distribution functions such that $\underline{F}_X = \underline{F}'_X$ and $\overline{F}_X = \overline{F}'_X$. In that case:*

1. *$\mathcal{F}_X \succeq_{FSD_1} \mathcal{F}_Y \Leftrightarrow \mathcal{F}'_X \succeq_{FSD_1} \mathcal{F}_Y$, and $\mathcal{F}_Y \succeq_{FSD_1} \mathcal{F}_X \Leftrightarrow \mathcal{F}_Y \succeq_{FSD_1} \mathcal{F}'_X$.*
2. *With respect to definitions $(FSD_2), \ldots, (FSD_6)$, if we assume that $\mathcal{F}_X \succ_{FSD_i} \mathcal{F}_Y$, then the possible scenarios for the relationship between $\mathcal{F}'_X$ and $\mathcal{F}_Y$ are summarised by the following table:*

	FSD_2	FSD_3	FSD_4	FSD_5	FSD_6
$\mathcal{F}'_X \succ_{FSD_i} \mathcal{F}_Y$	•	•	•	•	•
$\mathcal{F}_Y \succ_{FSD_i} \mathcal{F}'_X$		•	•		•
$\mathcal{F}'_X \equiv_{FSD_i} \mathcal{F}_Y$		•	•	•	•
$\mathcal{F}'_X, \mathcal{F}_Y$ *incomparable*	•	•	•		•

Proof Concerning definition (FSD_1), Proposition 4.19 assures that $\mathcal{F}_X \succeq_{FSD_1} \mathcal{F}_Y$ if and only if $\overline{F}_X = \overline{F}'_X \le \underline{F}_Y$, and using the same result this is equivalent to $\mathcal{F}'_X \succeq_{FSD_1} \mathcal{F}_Y$. The same result shows that $\mathcal{F}_Y \succeq_{FSD_1} \mathcal{F}_X$ if and only if $\overline{F}_Y \le \underline{F}_X = \underline{F}'_X$, and this is again equivalent to $\mathcal{F}_Y \succeq_{FSD} \mathcal{F}'_X$.
Let us prove that $\mathcal{F}'_X \succ_{FSD_2} \mathcal{F}_Y$ and $\mathcal{F}_Y \succ_{FSD_2} \mathcal{F}'_X$ are incompatible. If $\mathcal{F}_X \succ_{FSD_2} \mathcal{F}_Y$, then $\mathcal{F}_Y \not\succeq_{FSD_2} \mathcal{F}_X$. This means that for every $F_2 \in \mathcal{F}_Y$ there exist $F_1 \in \mathcal{F}_X$ and x_0 such that $F_1(x_0) < F_2(x_0)$. As a consequence,

$$\inf_{F_1 \in \mathcal{F}'_X} F_1(x_0) = \underline{F}'_X(x_0) = \underline{F}_X(x_0) \le F_1(x_0) < F_2(x_0),$$

whence for every $F_2 \in \mathcal{F}_Y$ there is some $F_1 \in \mathcal{F}'_X$ such that $F_1(x_0) < F_2(x_0)$, and consequently $F_2 \not\le F_1$. This means that $\mathcal{F}_Y \not\succeq_{FSD_2} \mathcal{F}'_X$, and therefore we cannot have $\mathcal{F}_Y \succ_{FSD_2} \mathcal{F}'_X$.

Let us show next that $\mathcal{F}_X \succeq_{FSD_5} \mathcal{F}_Y$ implies that $\mathcal{F}'_X \succeq_{FSD_5} \mathcal{F}_Y$. If $\mathcal{F}_X \succeq_{FSD_5} \mathcal{F}_Y$, there is $F_2 \in \mathcal{F}_Y$ such that $F_1 \le F_2$ for every $F_1 \in \mathcal{F}_X$. Whence, $\overline{F}_X \le F_2$, and therefore $\overline{F}'_X \le F_2$, which implies that also $F_1 \le F_2$ for every $F_1 \in \mathcal{F}'_X$. Hence, $\mathcal{F}'_X \succeq_{FSD_5} \mathcal{F}_Y$. ∎

Next example shows that the other scenarios are possible.

Example 4.41 *Let us consider sets of cumulative distribution functions F_X, F_X and F_Y that satisfies $\underline{F}_X = \underline{F}_X$ and $\overline{F}_X = \overline{F}_X$, and we are going to see that the scenarios given in Proposition 4.40 are possible.*

- *It is obvious that we can find some examples where* $F_X \; {}_{FSD_i} \; F_Y$ *for* $i = 2, \dots, 6$ *and* $F_X \; {}_{FSD_i} \; F_Y$. *To see it, it is enough to consider* $F_X = F_X$.
- *Let us show that* $F_X \; {}_{FSD_3} \; F_Y$ *and* $F_Y \; {}_{FSD_3} \; F_X$ *can hold simultaneously. Consider the sets:*

$$
\begin{aligned}
F_X &= \{F_{(0.5,0.5),(0,0.5)}, F_{(0.5,0.5),(0.5,1)}\},\\
F_X &= \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\},\\
F_Y &= \{F_{(0.75,0.25),(0.5,1)}, F_{(0.25,0.25,0.5),(0,0.5,1)}\}.
\end{aligned}
$$

It holds that $\underline{F}_X = \underline{F}_X$ *and* $\overline{F}_X = \overline{F}_X$. *Moreover it holds that* $F_X \; {}_{FSD_3} \; F_Y$ *since*

$$F_{(0.5,0.5),(0.5,1)} \; {}_{FSD} \; F_{(0.75,0.25),(0.5,1)}, F_{(0.25,0.25,0.5),(0,0.5,1)},$$

but for $F_{(0.5,0.5),(0.5,1)}$ *there is no distribution function in* F_Y *smaller than or equal to* $F_{(0.5,0.5),(0.5,1)}$. *Similarly,* $F_Y \; {}_{FSD_3} \; F_X$, *since*

$$
\begin{aligned}
&F_{(0.75,0.25),(0.5,1)} \; {}_{FSD} \; F_{1,0.5} \text{ and}\\
&F_{(0.25,0.25,0.5),(0,0.5,1)} \; {}_{FSD} \; F_{(0.5,0.5),(0,1)}.
\end{aligned}
$$

However, $F_{1,0.5}, F_{(0.5,0.5),(0,1)} \; {}_{FSD} \; F_{(0.25,0.25,0.5),(0,0.5,1)}$.

- *We now prove that the same can happen with Definition* (FSD_6). *Let us consider*

$$F_Y = \{F_{(0.25,0.75),(0,0.5)}, F_{(0.5,0.25,0.25),(0,0.5,1)}\}.$$

Then it holds that $F_X \; {}_{FSD_6} \; F_Y$ *and* $F_Y \; {}_{FSD_6} \; F_X$. *To check that* $F_X \; {}_{FSD_6} \; F_Y$ *it suffices to see that:*

$$
\begin{aligned}
&F_{1,0.5} \; {}_{FSD} \; F_{(0.25,0.75),(0,0.5)} \text{ and that}\\
&F_{(0.5,0.5),(0,1)} \; {}_{FSD} \; F_{(0.5,0.25,0.25),(0,0.5,1)},
\end{aligned}
$$

but $F_{(0.25,0.75),(0,0.5)} \; {}_{FSD} \; F_{1,0.5}, F_{(0.5,0.5),(0,1)}$. *To check that* $F_Y \; {}_{FSD_6} \; F_X$ *it suffices to see that*

$$F_{(0.25,0.75),(0,0.5)}, F_{(0.5,0.25,0.25),(0,0.5,1)} \; {}_{FSD} \; F_{(0.5,0.5),(0,0.5)}$$

but $F_{(0.5,0.5),(0,0.5)}$ *is not stochastically dominated by none of the distribution in* F_Y.

- *Next we prove that it is possible that* $F_X \; {}_{FSD_4} \; F_Y$ *and* $F_Y \; {}_{FSD_4} \; F_X$. *For this aim, we consider:*

$$
\begin{aligned}
F_X &= \{F_{(0.25,0.25,0.5),(0,0.5,1)}, F_{1,0.5}, F_{(0.5,0.5),(0,1)}\},\\
F_Y &= \{F_{(0.25,0.5,0.25),(0,0.5,1)}, F_{(0.4,0.2,0.4),(0,0.5,1)}\} \text{ and}\\
F_X &= \{F_{(0.25,0.75),(0,0.5)}, F_{1,0.5}, F_{(0.5,0.5),(0,1)}\}.
\end{aligned}
$$

It holds that $\underline{F}_X = \underline{F}_{X'}$ *and* $\overline{F}_X = \overline{F}_{X'}$. *Also*

$$F_{(0.25,0.25,0.5),(0,0.5,1)} \succ_{FSD} F_{(0.25,0.5,0.25),(0,0.5,1)},$$

but no distribution in $\mathcal{F}_Y$ *is dominated by a distribution function in* $\mathcal{F}_X$. *Whence* $\mathcal{F}_X \succ_{FSD_4} \mathcal{F}_Y$. *On the other hand,*

$$\begin{aligned}
&F_{(0.25,0.5,0.25),(0,0.5,1)} \succ_{FSD} F_{(0.25,0.75),(0,0.5)}, \text{ but}\\
&F_{(0.25,0.75),(0,0.5)} \succ_{FSD} F_{(0.25,0.5,0.25),(0,0.5,1)}, F_{(0.4,0.2,0.4),(0,0.5,1)}\\
&F_{1,0.5} \succ_{FSD} F_{(0.25,0.5,0.25),(0,0.5,1)}, F_{(0.4,0.2,0.4),(0,0.5,1)},\\
&F_{(0.5,0.5),(0,1)} \succ_{FSD} F_{(0.25,0.5,0.25),(0,0.5,1)}, F_{(0.4,0.2,0.4),(0,0.5,1)},
\end{aligned}$$

so $\mathcal{F}_Y \succ_{FSD_4} \mathcal{F}_{X'}$.

- *Let us now show that* $\mathcal{F}_X$ *may strictly dominate* $\mathcal{F}_Y$ *while* $\mathcal{F}_{X'}$ *and* $\mathcal{F}_Y$ *are indifferent when we consider definition* (FSD_i) *for* $i = 3, 4, 6$. *For this aim consider* $\mathcal{F}_X, \mathcal{F}_Y$ *associated with the same p-box and such that* $\mathcal{F}_X \succ_{FSD_i} \mathcal{F}_Y$ *for* $i = 3, \ldots, 6$, *as in Example 4.38, and let* $\mathcal{F}_{X'} = \mathcal{F}_Y$.
- *To see that* $\mathcal{F}_X \equiv_{FSD_5} \mathcal{F}_Y$ *and* $\mathcal{F}_{X'} \succ_{FSD_5} \mathcal{F}_Y$, *it is enough to consider the sets* $\mathcal{F}_X = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\}$, $\mathcal{F}_Y = \{F_{(0.5,0.5),(0,0.5)}, F_{(0.5,0.5),(0.5,1)}\}$ *and* $\mathcal{F}_{X'} = \mathcal{F}_Y$.
- *For* $\mathcal{F}_X \succ_{FSD_i} \mathcal{F}_Y$ *while* $\mathcal{F}_{X'}, \mathcal{F}_Y$ *are* (FSD_i) *incomparable for* $i = 2, 3, 4$, *take*

$$\begin{aligned}
\mathcal{F}_X &= \{F_{(0.5,0.5),(0.5,1)}, F_{(0.5,0.5),(0,0.5)}\},\\
\mathcal{F}_Y &= \{F\}, \text{ and}\\
\mathcal{F}_{X'} &= \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\}.
\end{aligned}$$

- *For* $\mathcal{F}_X \succ_{FSD_6} \mathcal{F}_Y$ *while* $\mathcal{F}_{X'}, \mathcal{F}_Y$ *are* (FSD_6) *incomparable, take*

$$\begin{aligned}
\mathcal{F}_X &= \{F_{(\frac{1}{n},1-\frac{2}{n},\frac{1}{n}),(0,0.5,1)}, F_{(\frac{1}{2}-\frac{1}{n},\frac{2}{n},\frac{1}{2}-\frac{1}{n}),(0,0.5,1)} \mid n \geq 3\},\\
\mathcal{F}_{X'} &= \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\},\\
\mathcal{F}_Y &= \{F_{(0.5-\frac{1}{n},0.5,\frac{1}{n}),(0,0.5,1)}, F \mid n \geq 3\}.
\end{aligned}$$

Remark 4.42 *Note that, under the conditions of the previous proposition, if we assume in addition that* $\underline{F}_X, \overline{F}_X \in \mathcal{F}_X \cap \mathcal{F}_{X'}$ *and that* $\underline{F}_Y, \overline{F}_Y \in \mathcal{F}_Y$, *then we deduce from Corollary 4.22 that* $\mathcal{F}_X \succ_{FSD_i} \mathcal{F}_Y \Leftrightarrow \mathcal{F}_{X'} \succ_{FSD_i} \mathcal{F}_Y$, *for* $i = 1, \ldots, 6$.

σ-additive VS finitely additive distribution functions

Although in this work we are focusing on sets of distribution functions associated with σ-additive probability measures, it is not uncommon to encounter situations where our imprecise information is given by means of sets of *finitely* additive probabilities: this is the case of the models of coherent lower and upper previsions in [205], and in particular of almost all mo dels of non-additive measures considered in the literature [126];in this sense the y are easier to handle than sets of σ-additive probability measures, which do not have an easy characterisation in terms of their lower and upp er envelopes, as showed in [102].

A finitely additive probability measure induces a finitely additive distribu tion function, and conversely, any finitely additive distribution function can be induced by a finitely additive probability measure [133]. As a consequence, given a p-b ox $(\underline{F}, \overline{F})$, the set of finitely additive probabilities compatible with this p-b ox induces the class of finitely additive dis tribution functions

$$\mathcal{F} := \{F \text{ finitelyadditive distributionfunction } : \underline{F} \leq F \leq \overline{F}\}. \tag{4.4}$$

In particu lar, b oth $\underline{F}, \overline{F}$ belong to $\mathcal{F}$. Takingthisintoaccount, ifwedefineconditions of sto chastic dominance analogous to those in Definition 4.1 for sets of finitely additive distribution fu nctions, it is not difficult to es tablish a characterisation similar to Corollary 4.22.

Lemma 4.43 *Let $\mathcal{F}_X, \mathcal{F}_Y$ betwo sets offinitely additivedistributionfunctions with associated p-boxes $(\underline{F}_X, \overline{F}_X)$, $(\underline{F}_Y, \overline{F}_Y)$. Assume $\underline{F}_X, \overline{F}_X$ $\mathcal{F}_X$ and $\underline{F}_Y, \overline{F}_Y$ $\mathcal{F}_Y$.*

1. $\mathcal{F}_X$ $_{FSD_1}$ $\mathcal{F}_Y$ $\overline{F}_X \leq \underline{F}_Y$.
2. $\mathcal{F}_X$ $_{FSD_2}$ $\mathcal{F}_Y$ $\mathcal{F}_X$ $_{FSD_3}$ $\mathcal{F}_Y$ $\underline{F}_X \leq \underline{F}_Y$.
3. $\mathcal{F}_X$ $_{FSD_4}$ $\mathcal{F}_Y$ $\underline{F}_X \leq \overline{F}_Y$.
4. $\mathcal{F}_X$ $_{FSD_5}$ $\mathcal{F}_Y$ $\mathcal{F}_X$ $_{FSD_6}$ $\mathcal{F}_Y$ $\overline{F}_X \leq \overline{F}_Y$.

Pro of The proof is analogous to the one for Corollary 4.22. ∎

We deduce in particular that under the ab ove conditions definitions (FSD_2) and (FSD_3) are equivalent, and the same applies to (FSD_5) and (FSD_6). Note that, although in this result we are using that the lower and upp er distribution functions of the p-b ox b elong to the asso ciated set of finitely additive distribution functions, this isnot necessary for the first statement.

In this section, we are going to investigate the relationship b etween the res ults we have obtained for sets of σ-additive probability measures and those we would obtain for finitely additive ones. Let $\mathcal{P}_X, \mathcal{P}_Y$ be two sets of σ-additive probability measures, and let $\mathcal{F}_X, \mathcal{F}_Y$ be their asso ciated sets of distribution functions. Thesesets ofdistribution functions determine p-b oxes $(\underline{F}_X, \overline{F}_X)$, $(\underline{F}_Y, \overline{F}_Y)$. Let $\mathcal{F}_X, \mathcal{F}_Y$ be two sets of finitely additive distribution functions asso ciated with the p-b oxes $(\underline{F}_X, \overline{F}_X)$, $(\underline{F}_Y, \overline{F}_X)$.

When the lower and upp er distribution functions of the asso ciated p-b ox b elong toour set of cumulative distribution functions, we can easily show that the sto chastic dominance holds under the same conditions regardless of whether we work with finitely or σ-additive probability measures:

Corollary 4.44 *Let us consider two sets of cumulative distribution functions $\mathcal{F}_X$ and $\mathcal{F}_Y$ with associated p-boxes $(\underline{F}_X,\overline{F}_X)$, $(\underline{F}_Y,\overline{F}_Y)$, and let $\mathcal{F}_X,\mathcal{F}_Y$ be the sets of finitely additive distribution functions associated with these p-boxes. If $\underline{F}_X,\overline{F}_X \in \mathcal{F}_X$ and $\underline{F}_Y,\overline{F}_Y \in \mathcal{F}_Y$, it holds that:*

$$\mathcal{F}_X \succeq_{FSD_i} \mathcal{F}_Y \Leftrightarrow \mathcal{F}_X \succeq_{FSD_i} \mathcal{F}_Y,$$

for $i = 1,\ldots,6$.

Proof The result is an immediate consequence of Corollary 4.22 and Lemma 4.43. ■

However, when the lower and the upper distribution functions induced by $\mathcal{F}_X$ and $\mathcal{F}_Y$ do not belong to these sets, the equivalence no longer holds. We can nonetheless establish the following result:

Proposition 4.45 *Let us consider two sets of cumulative distribution functions $\mathcal{F}_X$ and $\mathcal{F}_Y$, and two sets of finite distribution functions $\mathcal{F}_X$ and $\mathcal{F}_Y$ such that $\mathcal{F}_X,\mathcal{F}_X$ induce the same p-box $(\underline{F}_X,\overline{F}_X)$ and $\mathcal{F}_Y,\mathcal{F}_Y$ induce the same p-box $(\underline{F}_Y,\overline{F}_Y)$. Then:*

1. $\mathcal{F}_X \succeq_{FSD_1} \mathcal{F}_Y \Leftrightarrow \mathcal{F}_X \succeq_{FSD_1} \mathcal{F}_Y$.
2. *The relationship $\mathcal{F}_X \succeq_{FSD_i} \mathcal{F}_Y$ does not have any implication in general on the relationship between $\mathcal{F}_X$ and $\mathcal{F}_Y$ with respect to (FSD_i), for* $i = 2,3,4,5,6$.

Proof

1. From Proposition 4.19, we know that $\mathcal{F}_X \succeq_{FSD_1} \mathcal{F}_Y \Leftrightarrow \overline{F}_X \le \underline{F}_Y$. The same proof allows to show the equivalence with $\mathcal{F}_X \succeq_{FSD_1} \mathcal{F}_Y$.
2. If we apply Proposition 4.40 with $\mathcal{F}_Y = \mathcal{F}_Y$, we see that all we need to prove is that $\mathcal{F}_X \succeq_{FSD_i} \mathcal{F}_Y$ is compatible with $\mathcal{F}_Y \succeq_{FSD_i} \mathcal{F}_X$ for $i = 2,5$, with $\mathcal{F}_X \equiv_{FSD_i} \mathcal{F}_Y$ for $i = 2$ and with $\mathcal{F}_X,\mathcal{F}_Y$ incomparable with respect to (FSD_5). ■

Next we give examples of all the possibilities in the previous result.

Example 4.46 *Let us show that, given two sets $\mathcal{F}_X,\mathcal{F}_X$, with $(\underline{F}_X,\overline{F}_X) = (\underline{F}_X,\overline{F}_X)$, and $\mathcal{F}_Y,\mathcal{F}_Y$, with $(\underline{F}_Y,\overline{F}_Y) = (\underline{F}_Y,\overline{F}_Y)$, a number of preference scenarios are possible (the other possible scenarios have already been established in the proof).*

We begin by showing that wemayhave $F_X \quad {}_{FSD_2}\, F_Y$ *and* $F_Y \quad {}_{FSD_2}\, F_X$. *To see this, consider* F_X, F_Y *defined by:*

$$F_X = \{F_{(0.5,0.5),(0,0.5)}, F_{(0.5,0.5),(0.5,1)}\} \text{ and}$$
$$F_Y = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\}.$$

They are associatedwith the same p-box and satisfy $F_X \quad {}_{FSD_2}\, F_Y$. *Wealso consider* $F_X = F_Y, F_Y = F_X$. *Asimilarreasoningshowsthatwemay have* $F_X \quad {}_{FSD_5}\, F_Y$ *while* $F_Y \quad {}_{FSD_5}\, F_Y$.

Next, weshow that we may have $F_X \quad {}_{FSD_2}\, F_Y$ *and* $F_X \equiv_{FSD_2} F_Y$. *Let*

$$F_X = F_X = F_Y = \{F_{(0.5,0.5)(0,0.5)}, F_{(0.5,0.5),(0.5,1)}\} \text{ and}$$
$$F_Y = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\}.$$

It can be easily seen that $F_X \quad {}_{FSD_2}\, F_Y$ *and that* F_X, F_Y *inducethe same p-box. Since* $F_{(0.5,0.5),(0.5,1)} \quad F_X \cap F_Y$ *satisfies that* $F_{(0.5,0.5),(0.5,1)} \leq F_{(0.5,0.5),(0,0.5)}$, *we deduce that* $F_X \equiv_{FSD_2} F_Y$.

To conclude, wegive anexamplewhere $F_X \quad {}_{FSD_5}\, F_Y$ *while* F_X, F_Y *are incomparable wit h respect to* (FSD_5). *Consider the sets cumulative distribution functions*

$$F_X = F_X = \{F_{(\frac{1}{n},1-\frac{2}{n},\frac{1}{n}),(0,0.5,1)} \mid n \geq 3\},$$
$$F_Y = \{F_{(0.5,0.5),(0,0.5)}, F_{(0.5,0.5),(0.5,1)}\} \text{ and}$$
$$F_Y = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\}.$$

Then $F_X \quad {}_{FSD_5}\, F_Y$ *because* $F_{(\frac{1}{n},1-\frac{2}{n},\frac{1}{n}),(0,0.5,1)} \leq F_{(0.5,0.5),(0,0.5)}$ *for every* $n \geq 3$. *On the other hand,* F_X *and* F_Y *are incomparable with respectto* (FSD_5).

It is known that any finitely additive cumulative distribution function F can be approximated bya σ-additive cumulative distribution function F : its right-continuous approximation, given by

$$F(x) = \inf_{y>x} F(y) \quad x < 1, \quad F(1) = 1. \tag{4.5}$$

Hence, toany set F of finitely additive cumulative distribution functions we can asso ciate aset F of σ-additive cumulative distribution functions, defi ned by $F := \{F : F \quad F\}$, and where F is given byEquation (4.5). However, both sets do not mo del the same preferences, as wecan see from the followingresult:

Prop osition 4.47 *Let* F *be a set of finitely additive cu mulative distribution functions, and let* F *bethe setof their σ-additive approximations. Therelationships between* F *and* F *are summarised in the fol lowing table:*

	FSD_1	FSD_2	FSD_3	FSD_4	FSD_5	FSD_6
$F \quad _{FSD_i} F$	•	•	•	•	•	•
$F \quad _{FSD_i} F$						
$F \equiv_{FSD_i} F$	•	•	•	•	•	•
F, F incomparable	•	•			•	

Proof From Equation (4.5), $F \leq F$ for any $F \quad F$, whence $F \quad _{FSD_i} F$, for $i = 3, 4, 6$. We deduce from Proposition 4.3 that we cannot have $F \quad _{FSD_i} F$ for $i = 1, \ldots, 6$. ∎

Next example shows that the remaining scenarios are possible.

Example 4.48 *If F_1 is a σ-additive distribution function and we take $F = \{F_1\}$, we obtain $F = F = \{F_1\}$, and $F \equiv_{FSD_i} F$ for $i = 1, \ldots, 6$.*

On the other hand, if $F_1 = I_{(0.5,1]}$ and $F = \{F_1\}$, we obtain that $F_1 = I_{[0.5,1]}$, whence $F_1 < F_1$ and as a consequence $F \quad _{FSD_i} F$ for $i = 1, \ldots, 6$.

Finally, if $F = \{I_{[x,1]} : x \quad (0,1)\}$, we obtain that $F = F$ and F is incomparable with itself with respect to conditions (FSD_i) for $i = 1, 2, 5$.

Convergence of p-boxes

It is well-known that a distribution function can be seen as the limit of the empirical distribution function that we derive from a sample, as we increase the sample size. Something similar applies when we consider a set of distribution functions: it was proven in [136] that any p-box on the unit interval is the limit of a sequence of p-boxes $(\underline{F}_n, \overline{F}_n)_n$ that are *discrete*, in the sense that for every n both $\underline{F}_n$ and $\overline{F}_n$ have a finite number of discontinuity points.

If for two given p-boxes $(\underline{F}_X, \overline{F}_X)$, $(\underline{F}_Y, \overline{F}_Y)$ we consider respective approximating sequences $(\underline{F}_{X,n}, \overline{F}_{X,n})_n$, $(\underline{F}_{Y,n}, \overline{F}_{Y,n})_n$, in the sense that

$$\lim_n \underline{F}_{X,n} = \underline{F}_X, \ \lim_n \overline{F}_{X,n} = \overline{F}_X, \ \lim_n \underline{F}_{Y,n} = \underline{F}_Y, \ \lim_n \overline{F}_{Y,n} = \overline{F}_Y,$$

we wonder if it is possible to say something about the preferences between $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$ by comparing for each n the discrete p-boxes $(\underline{F}_{X,n}, \overline{F}_{X,n})$ and $(\underline{F}_{Y,n}, \overline{F}_{Y,n})$. This is what we set out to do in this section. We shall be even more general, by considering sets of distribution functions whose associated p-boxes converge to some limit.

Proposition 4.49 *Let $(F_{X,n})_n, (F_{Y,n})_n$ be two sequences of sets of distribution functions and let us denote their associated sequences of p-boxes by $(\underline{F}_{X,n}, \overline{F}_{X,n})$ and $(\underline{F}_{Y,n}, \overline{F}_{Y,n})$ for $n \quad \mathbb{N}$. Let F_X, F_Y be two sets of cumulative distribution functions with associated* *p-boxes* $(\underline{F}_X, \overline{F}_X)$ *and* $(\underline{F}_Y, \overline{F}_Y)$. *Letus assume that:*

$$\overline{F}_{X,n} \xrightarrow{n} \overline{F}_X \qquad \underline{F}_{X,n} \xrightarrow{n} \underline{F}_X$$
$$\overline{F}_{Y,n} \xrightarrow{n} \overline{F}_Y \qquad \underline{F}_{Y,n} \xrightarrow{n} \underline{F}_Y$$

and that $\underline{F}_X, \overline{F}_X \quad \mathcal{F}_X$ *and* $\underline{F}_Y, \overline{F}_Y \quad \mathcal{F}_Y$. *Then,* $\mathcal{F}_{X,n} \quad {}_{FSD_i} \ \mathcal{F}_{Y,n} \quad n$, *implies that* $\mathcal{F}_X \quad {}_{FSD_i} \ \mathcal{F}_Y$, *for* $i = 1, \ldots, 6$.

Pro of The result immediately follows from Propositions 4.3 and 4.19 and Corollary 4.22. ■

It follows from the pro of ab ove that the assumption that the upp er and lower distribution functions b elong to the corresp onding sets of distribution is not necessary for the implication with resp ect to (FSD_1); however, it is nece ssary for the other definitions, as we cansee in the next example.

Example 4.50 *Let us consider the fol lowing sets of cumulative distribution functions:*

$$\mathcal{F}_X = \{F_{1,0.5}, F_{(0.5,0.5),(0,1)}\}.$$
$$\mathcal{F}_{X,n} = \{F_{(0.5,0.5),(0,0.5)}, F_{(0.5,0.5),(0.5,1)}\}.$$
$$\mathcal{F}_Y = \mathcal{F}_{Y,n} = \{F\}.$$

$\mathcal{F}_X$ *and* $\mathcal{F}_Y$ *areincomparablewithrespect to* (FSD_4), *and consequently withrespect to* (FSD_i), *for* $i = 1, \ldots, 6$. *However,* $\mathcal{F}_{X,n} \quad {}_{FSD_i} \ \mathcal{F}_{Y,n}$ *for* $i = 2, 3, 4$ *and* $\mathcal{F}_{Y,n} \quad {}_{FSD_i} \ \mathcal{F}_{X,n}$ *for* $i = 4, 5, 6$.

Sto chastic dominance between p ossibility measures

So far, we have explored the extension of the notion of sto chastic dominance towards sets of probability measu re s, and we have showed that in some cases it is equivalent to compare the p-b oxes they determine.In this section, we are going to use sto chastic dominance to compare p ossibility measures asso ciated with *continuous* distribution functions. Recall that, from Definition2.41, a possibility measure Π is a supremum preserving function $\Pi: \mathcal{P}([0,1]) \to [0,1]$ andit ischaracterised byits restrictionto events π, called possibility distribution. Given two possibility measures Π_1 and Π_2, we can consider their asso ciated credal sets, given by Equation(2.19):

$$\mathcal{M}(\Pi_1) := \{P \text{ probability} : P(A) \leq \Pi_1(A) \quad A\}, \text{ and}$$
$$\mathcal{M}(\Pi_2) := \{P \text{ probability} : P(A) \leq \Pi_2(A) \quad A\}.$$

From these credal sets, we can also consider their asso ciated sets of distribution functions and their associated p-b oxes, given in Equation (2.20) by

$$\overline{F}_1(x) = \sup_{y \leq x} \pi_1(y), \quad \underline{F}_1(x) = 1 - \sup_{y > x} \pi_1(y),$$
$$\overline{F}_2(x) = \sup_{y \leq x} \pi_2(y), \quad \underline{F}_2(x) = 1 - \sup_{y > x} \pi_2(y).$$

When considering possibility measures associated with continuous distribution functions, both the lower and the upper distribution functions belong to the set of distribution functions associated with the possibility measures:

Lemma 4.51 *Let Π be a possibility measure associated with a continuous possibility distribution on [0,1]. Then, there exist probability measures $P_1, P_2 \in M(\Pi)$ whose associated distribution functions are $F_{P_1} = \overline{F}, F_{P_2} = \underline{F}$.*

Proof Let us consider the probability space $([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$, where $\beta_{[0,1]}$ denotes the Borel σ-field and $\lambda_{[0,1]}$ the Lebesgue measure, and let $\Gamma : [0,1] \to P([0,1])$ be the random set given by $\Gamma(\alpha) = \{x : \pi(x) \geq \alpha\} = \pi^{-1}([\alpha, 1])$. Then it was proved in [84] that Π is the upper probability of Γ.

Let us consider the mappings $U_1, U_2 : [0,1] \to [0,1]$ given by $U_1(\alpha) = \min \Gamma(\alpha)$, $U_2(\alpha) = \max \Gamma(\alpha)$. Since we are assuming that π is a continuous mapping, the set $\pi^{-1}([\alpha, 1]) = \Gamma(\alpha)$ has a maximum and a minimum value for every $\alpha \in [0,1]$, so U_1, U_2 are well-defined. It also follows that U_1, U_2 are measurable mappings, and as a consequence the probability measures they induce P_{U_1}, P_{U_2} belong to the set $M(\Pi)$. Their associated distribution functions are:

$$
\begin{aligned}
F_{U_1}(x) &= P_{U_1}([0,x]) = \lambda_{[0,1]}(U_1^{-1}([0,x])) = \lambda_{[0,1]}(\{\alpha : \min \Gamma(\alpha) \leq x\}) \\
&= \lambda_{[0,1]}(\{\alpha : \exists y \leq x : \pi(y) \geq \alpha\}) = \lambda_{[0,1]}(\{\alpha : \Pi[0,x] \geq \alpha\}) \\
&= \Pi([0,x]) = \overline{F}(x),
\end{aligned}
$$

where the fifth equality follows from the continuity of $\lambda_{[0,1]}$, and similarly

$$
\begin{aligned}
F_{U_2}(x) &= P_{U_2}([0,x]) = \lambda_{[0,1]}(U_2^{-1}([0,x])) = \lambda_{[0,1]}(\{\alpha : \max \Gamma(\alpha) \leq x\}) \\
&= \lambda_{[0,1]}(\{\alpha : \pi(y) < \alpha \ \forall y > x\}) = \lambda_{[0,1]}(\{\alpha : \Pi(x,1] \leq \alpha\}) \\
&= 1 - \Pi((x,1]) = \underline{F}(x),
\end{aligned}
$$

again using the continuity of $\lambda_{[0,1]}$. Hence, $\overline{F}, \underline{F}$ belong to the set of distribution functions induced by $M(\Pi)$. ■

As a consequence, if we consider two possibility measures Π_1, Π_2 with continuous possibility distributions π_1, π_2, the lower and upper distribution functions of their respective p-boxes belong to the sets F_1, F_2. Hence, we can apply Proposition 4.21 and conclude that $F_1 \succeq_{FSD_2} F_2 \Leftrightarrow F_1 \succeq_{FSD_3} F_2$ and $F_1 \succeq_{FSD_5} F_2 \Leftrightarrow F_1 \succeq_{FSD_6} F_2$. Moreover, we can use Corollary 4.22 and conclude that:

$$
\begin{aligned}
F_1 \succeq_{FSD_1} F_2 &\Leftrightarrow \overline{F}_1 \leq \underline{F}_2 \\
F_1 \succeq_{FSD_2} F_2 &\Leftrightarrow \underline{F}_1 \leq \underline{F}_2 \\
F_1 \succeq_{FSD_4} F_2 &\Leftrightarrow \underline{F}_1 \leq \overline{F}_2 \\
F_1 \succeq_{FSD_5} F_2 &\Leftrightarrow \overline{F}_1 \leq \overline{F}_2.
\end{aligned}
$$

The following proposition gives a sufficient condition for each of these relationships.

Prop osition 4.52 *Let F_1, F_2 bethesets of distribution functions associated with the possibility measures Π_1, Π_2.*

1. $\Pi_1 \leq N_2$ $F_1 \;{}_{FSD_1}\; F_2$.
2. $\Pi_2 \leq \Pi_1$ $F_1 \;{}_{FSD_2}\; F_2, F_1 \;{}_{FSD_3}\; F_2$
3. $M(\Pi_1) \cap M(\Pi_2) =$ $F_1 \;{}_{FSD_4}\; F_2$.
4. $N_2 \leq N_1$ $F_1 \;{}_{FSD_5}\; F_2, F_1 \;{}_{FSD_6}\; F_2$.

Pro of

1. Note that $\overline{F}_1 \leq \underline{F}_2$ if and only if $\sup_{y \leq x} \pi_1(y) \leq 1 - \sup_{y>x} \pi_2(y)$ for every x, or, equivalently, if and onlyif $\Pi_1([0, x]) \leq 1 - \Pi_2((x, 1]) = N_2([0, x])$ for every x. Then, if $\Pi_1(A) \leq N_2(A)$ for any A, in particu lar the inequality holds for the sets $[0, x]$ and therefore $\overline{F}_1 \leq \underline{F}_2$.

2. Similarly, $\underline{F}_1 \leq \underline{F}_2$ if and only if $1 - \sup_{y \leq x} \pi_1(y) \leq 1 - \sup_{y>x} \pi_2(y)$ for every x, or, equivalently, if and only if $\Pi_2((x, 1]) \leq \Pi_1((x, 1])$ for every x. Then, if $\Pi_2(A) \leq \Pi_1(A)$ for any A, in particular the inequality holds for the sets $(x, 1]$, and therefore $\underline{F}_1 \leq \underline{F}_2$.

3. For the fourth condition of sto chastic dominance, note that $\underline{F}_1 \leq \overline{F}_2$ ifand only if $1 - \sup_{y>x} \pi_1(y) \leq \sup_{y \leq x} \pi_2(y)$ for every x, or, equivalently, if and only if $1 \leq \Pi_1((x, 1]) + \Pi_2([0, x])$ for every x. As a consequence, ifthere is a probability P $M(\Pi_1) \cap M(\Pi_2)$,

$$1 = P((x, 1]) + P([0, x]) \leq \Pi_1((x, 1]) + \Pi_2([0, x]),$$

   whence $F_1 \;{}_{FSD_4}\; F_2$.

4. Finally, note that $\overline{F}_1 \leq \overline{F}_2$ if and only if $\sup_{y \leq x} \pi_1(y) \leq \sup_{y \leq x} \pi_2(y)$ for every x, or, equivalently, if and onlyif $\Pi_1([0, x]) \leq \Pi_2([0, x])$ for every x. Hence, if $\Pi_1 \leq \Pi_2$ (or, equivalently, if $N_2 \leq N_1$) we have that $F_1 \;{}_{FSD_5}\; F_2$ and $F_1 \;{}_{FSD_6}\; F_2$. ∎

However, none of the ab ove conditions is necessary, as we show in the next example.

Example 4.53 *1. First of al l, let usseethat $F_X \;{}_{FSD_1}\; F_Y$ $\Pi_X \leq N_Y$. For this aim, let π_X, π_Y be givenby*

$$\pi_X(x) = \begin{cases} 0 & \text{if } x \leq 0.5 \\ 2x - 1 & \text{otherwise,} \end{cases} \quad \text{and } \pi_Y(x) = \begin{cases} 1 & \text{if } x \leq 0.5 \\ 2 - 2x & \text{otherwise.} \end{cases}$$

Then for every x $[0, 1]$ *it holds that* $\Pi_X([0, x]) + \Pi_Y((x, 1]) \leq 1$*: this holds trivial ly for* $x \leq 0.5$ *because in that case* $\Pi_X([0, x]) = 0$. *For* $x > 0.5$*, we have that*

$$\Pi_X([0, x]) + \Pi_Y((x, 1]) = 2x - 1 + 2 - 2x = 1.$$

Hence, F_X $_{FSD_1}$ F_Y*. However:*

$$\Pi_X([0.5, 1]) = 1 > N_Y([0.5, 1]) = 1 - \Pi_Y([0, 0.5)) = 1 - 1 = 0,$$

so the converse of the first implication does not hold.

2. *Now, weare goingto see that* F_X $_{FSD_2, FSD_3}$ F_Y $\Pi_Y \leq \Pi_X$*. Consider the possibility distributions* π_X, π_Y *given by*

$$\pi_X(x) = x, \quad \pi_Y(x) = 1 \quad x.$$

Then $\Pi_Y((x, 1]) = 1 = \Pi_X((x, 1])$ *for all* x*, whence* F_X $_{FSD_2}$ F_Y*. However,* $\Pi_X([0, 0.5]) = 0.5 < 1 = \Pi_Y([0, 0.5])$*, so* Π_Y Π_X*.*

3. *Now we are going to seethat* F_X $_{FSD_4}$ F_Y $M(\Pi_X) \cap M(\Pi_Y) =$ *. Let* π_X, π_Y *be given by*

$$\pi_X(x) = \begin{cases} 4x - 3 & \text{if } x \geq 0.75 \\ 0 & \text{otherwise.} \end{cases} \text{ and } \pi_Y(x) = \begin{cases} 1 - 4x & \text{if } x \leq 0.25 \\ 0 & \text{otherwise.} \end{cases}$$

Then for every x $[0, 1]$ *it holds that*

$$\Pi_X((x, 1]) + \Pi_Y([0, x]) \geq \Pi_Y([0, x]) = 1,$$

whence F_X $_{FSD_4}$ F_Y*. However, any probability* P *in* $M(\Pi_X) \cap M(\Pi_Y)$ *should satisfy*

$$P([0, 0.5]) \leq \Pi_X([0, 0.5]) = 0, P((0.5, 1]) \leq \Pi_Y((0.5, 1]) = 0.$$

Hence, $M(\Pi_X) \cap M(\Pi_Y) =$ *.*

4. *Final ly, weare goingtoseethat* F_X $_{FSD_5, FSD_6}$ F_Y $\Pi_X \leq \Pi_Y$*. Consider the possibility distributions* π_X, π_Y *given by*

$$\pi_X(x) = 1, \quad \pi_Y(x) = 1 - x, \quad x.$$

Then it holds that $\Pi_X([0, x]) \leq \Pi_Y([0, x])$ x*, whence* F_X $_{FSD_5}$ F_Y*. However,* $\Pi_X([0.5, 1]) = 1 > 0.5 = \Pi_Y([0.5, 1])$*, so* Π_X Π_Y*.*

An op en problem from this section would be to apply the notion of stochastic dominance to compare p ossibil ity measures whose distributions are not necessarily continuous.

P-boxes where one of the bounds is trivial

To conclude this section we investigate the case of p-boxes where one of the bounds is trivial. These have been related to possibility and maxitive measures in [199], and consequently they are in some sense related to the previous paragraph. We shall show that when the lower distribution function is trivial, then the second and third conditions, which are based on the comparison of this bound, always produce indifference.

Proposition 4.54 *Let us consider the p-boxes $\mathcal{F}_X = (\underline{F}_X, \overline{F}_X)$ and $\mathcal{F}_Y = (\underline{F}_Y, \overline{F}_Y)$. Let us assume that $\underline{F}_X = \underline{F}_Y = I_{\{1\}}$, $\underline{F}_X \neq \overline{F}_X$ and $\underline{F}_Y \neq \overline{F}_Y$. Then the possible relationships between $\mathcal{F}_X$ and $\mathcal{F}_Y$ are:*

	FSD_1	FSD_2	FSD_3	FSD_4	FSD_5	FSD_6
$\mathcal{F}_X \succ_{FSD_i} \mathcal{F}_Y$					•	•
$\mathcal{F}_Y \succ_{FSD_i} \mathcal{F}_X$					•	•
$\mathcal{F}_X \equiv_{FSD_i} \mathcal{F}_Y$		•	•	•	•	•
$\mathcal{F}_X, \mathcal{F}_Y$ *incomparable*	•				•	•

Proof

- Using Proposition 4.19 we know that $\mathcal{F}_X \succ_{FSD_1} \mathcal{F}_Y \Leftrightarrow \overline{F}_X \leq \underline{F}_Y$. However, this cannot happen since $\underline{F}_Y = I_{\{1\}}$ and the p-boxes are not trivial. Consequently, both sets are incomparable with respect to (FSD_1).
- Since $\underline{F}_X = \underline{F}_Y \in \mathcal{F}_X \cap \mathcal{F}_Y$, we deduce from Corollary 4.22 that $\mathcal{F}_X \equiv_{FSD_2} \mathcal{F}_Y$. Applying Proposition 4.3, we deduce that $\mathcal{F}_X \equiv_{FSD_3} \mathcal{F}_Y$ and $\mathcal{F}_X \equiv_{FSD_4} \mathcal{F}_Y$.
- On the other hand, it is easy to see that anything can happen for definitions (FSD_5) and (FSD_6), since these depend on the upper cumulative distribution functions of the p-boxes. ■

Similarly, when the upper distribution function is trivial, then the fifth and sixth conditions, which are based on the comparison of these bounds, always produce indifference.

Proposition 4.55 *Let us consider the p-boxes $\mathcal{F}_X = (\underline{F}_X, \overline{F}_X)$ and $\mathcal{F}_Y = (\underline{F}_Y, \overline{F}_Y)$. Let us assume that $\overline{F}_X = \overline{F}_Y = 1$, $\underline{F}_X < \overline{F}_X$ and $\underline{F}_Y < \overline{F}_Y$. Then the possible relationships between $\mathcal{F}_X$ and $\mathcal{F}_Y$ are:*

	FSD_1	FSD_2	FSD_3	FSD_4	FSD_5	FSD_6
$\mathcal{F}_X \succ_{FSD_i} \mathcal{F}_Y$		•	•			
$\mathcal{F}_Y \succ_{FSD_i} \mathcal{F}_X$		•	•			
$\mathcal{F}_X \equiv_{FSD_i} \mathcal{F}_Y$		•	•	•	•	•
$\mathcal{F}_X, \mathcal{F}_Y$ *incomparable*	•	•	•			

Proof This proof is analogous to the previous one. ∎

This case is related to the previous paragraph devoted to possibility measures: when the lower distribution function is trivial, the probability measures determined by the p-box are those dominated by the possibility measure that has $\overline{F}$ as a possibility distribution; however, a similar result does not hold for the case of $(\underline{F}, 1)$ in general, because we need $\underline{F}$ to be right-continuous.

0-1-valued p-boxes

Let us now focus on 0-1-valued p-boxes, by which we mean p-boxes where both the lower and upper cumulative distribution functions $\underline{F}, \overline{F}$ are 0-1-valued. As we shall see, the notions of stochastic dominance will be related to the orderings between the intervals of the real line determined by these 0-1-valued distribution functions. 0-1-valued p-boxes have also been related to possibility measures in [199].

Given a 0-1-valued distribution function F, we denote

$$x_F = \inf\{x \mid F(x) = 1\}.$$

Note that this infimum is a minimum when we consider distribution functions associated with σ-additive probability measures, but not necessarily for those associated with finitely additive probability measures.

Using this notation and Proposition 4.19, we can characterise the comparison of sets of 0–1 valued distribution functions:

Proposition 4.56 *Let $\mathcal{F}_X$ and $\mathcal{F}_Y$ be two sets of cumulative distribution functions, with associated p-boxes $(\underline{F}_X, \overline{F}_X)$, $(\underline{F}_Y, \overline{F}_Y)$.*

a) If $\underline{F}_X$, $\overline{F}_X$, $\underline{F}_Y$ and $\overline{F}_Y$ are 0-1-valued functions, then

1. $\mathcal{F}_X \quad \text{FSD}_1 \quad \mathcal{F}_Y \qquad x_{\overline{F}_X} \geq x_{\underline{F}_Y}$.
2. $\mathcal{F}_X \quad \text{FSD}_2 \quad \mathcal{F}_Y \qquad x_{\underline{F}_X} \geq x_{\underline{F}_Y}$.
3. $\mathcal{F}_X \quad \text{FSD}_3 \quad \mathcal{F}_Y \qquad x_{\underline{F}_X} \geq x_{\underline{F}_Y}$.
4. $\mathcal{F}_X \quad \text{FSD}_4 \quad \mathcal{F}_Y \qquad x_{\underline{F}_X} \geq x_{\overline{F}_Y}$.
5. $\mathcal{F}_X \quad \text{FSD}_5 \quad \mathcal{F}_Y \qquad x_{\overline{F}_X} \geq x_{\overline{F}_Y}$.
6. $\mathcal{F}_X \quad \text{FSD}_6 \quad \mathcal{F}_Y \qquad x_{\overline{F}_X} \geq x_{\overline{F}_Y}$.

Moreover, if $\underline{F}_X, \overline{F}_X \quad \mathcal{F}_X$ and $\underline{F}_Y, \overline{F}_Y \quad \mathcal{F}_Y$, the converses also hold.

b) If in particular $\mathcal{F}_X$ and $\mathcal{F}_Y$ are two sets of 0-1 cumulative distribution functions it also holds that

2. $x_{\underline{F}_X} > x_{\underline{F}_Y}$ $\mathcal{F}_X$ $_{FSD_2}$ $\mathcal{F}_Y$ $\mathcal{F}_X$ $_{FSD_2}$ $\mathcal{F}_Y$.
3. $x_{\underline{F}_X} > x_{\underline{F}_Y}$ $\mathcal{F}_X$ $_{FSD_3}$ $\mathcal{F}_Y$ $\mathcal{F}_X$ $_{FSD_3}$ $\mathcal{F}_Y$.
4. $x_{\underline{F}_X} > x_{\overline{F}_Y}$ $\mathcal{F}_X$ $_{FSD_4}$ $\mathcal{F}_Y$.
5. $x_{\overline{F}_X} > x_{\overline{F}_Y}$ $\mathcal{F}_X$ $_{FSD_5}$ $\mathcal{F}_Y$ $\mathcal{F}_X$ $_{FSD_5}$ $\mathcal{F}_Y$.
6. $x_{\overline{F}_X} > x_{\overline{F}_Y}$ $\mathcal{F}_X$ $_{FSD_6}$ $\mathcal{F}_Y$ $\mathcal{F}_X$ $_{FSD_6}$ $\mathcal{F}_Y$.

Pro of In order to prove the first item of this result it is enough to consider Proposition 4.19, andto note that, if F and G are two0-1 finitely additivedistribution functions then $F \leq G$ implies that $x_F \geq x_G$. In particular, if G is a cumulative distribution function, $F \leq G$ if and only if $x_F \geq x_G$, from whichwe deduce that $x_{\overline{F}_X} \geq x_{\underline{F}_Y}$ $\mathcal{F}_X$ $_{FSD_1}$ $\mathcal{F}_Y$.

Moreover, if $\underline{F}_X, \overline{F}_X$ $\mathcal{F}_X$ and $\underline{F}_Y, \overline{F}_Y$ $\mathcal{F}_Y$, these arecumulative distribution functions, andwecanuse that $F \leq G$ if andonly if $x_F \geq x_G$. Applying Corollary4.22 we deducethat in that case the converse implications also hold.

Let us consider the second part. Onthe one hand,itisobviousthat $\mathcal{F}_X$ $_{FSD_i}$ $\mathcal{F}_Y$ implies $\mathcal{F}_X$ $_{FSD_i}$ $\mathcal{F}_Y$ for $i = 2, 3, 5, 6$. Let us check the other implicati on s.

2. If $x_{\underline{F}_X} > x_{\underline{F}_Y}$, x_0 such that $x_{\underline{F}_X} > x_0 > x_{\underline{F}_Y}$. Then, since $x_0 > x_{\underline{F}_Y}$, $\underline{F}_Y(x_0) = 1$ and therefore $F_2(x_0) = 1$ F_2 $\mathcal{F}_Y$. Since $x_{\underline{F}_X} > x_0$, $\underline{F}_X(x_0) = 0$ and as we are considering only $0-1$ valued cumulative distribution functions, there issome F_1 $\mathcal{F}_X$ such that $F_1(x_0) = 0$. Thus,

$$F_1 \quad \mathcal{F}_X \text{ such that } F_1 \quad {}_{FSD} \; F_2 \quad F_2 \quad \mathcal{F}_Y.$$

Then, $\mathcal{F}_X$ $_{FSD_2}$ $\mathcal{F}_Y$ and $\mathcal{F}_Y$ $_{FSD_2}$ $\mathcal{F}_X$. On the otherhand, if $\mathcal{F}_X$ $_{FSD_2}$ $\mathcal{F}_Y$, Prop osition 4.19 implies that $\underline{F}_X$ $_{FSD}$ $\underline{F}_Y$, and moreover the preference must be strict (otherwise both sets would be indifferent). Then, $x_{\underline{F}_X} > x_{\underline{F}_Y}$.

3. On the one hand, the direct implication follows from the previous item and Prop osition 4.15. Ontheotherhand, if $\mathcal{F}_X$ $_{FSD_3}$ $\mathcal{F}_Y$, by Prop osition 4.19 we know that $\underline{F}_X$ $_{FSD}$ $\underline{F}_Y$, andthe preferenceisinfact strict(otherwisethe sets $\mathcal{F}_X$ and $\mathcal{F}_Y$ would be indifferent). Then, following the sam e steps than in the previous item we conclude that $x_{\underline{F}_X} > x_{\underline{F}_Y}$.

4. If $x_{\underline{F}_X} > x_{\overline{F}_Y}$, x_0 such that $x_{\underline{F}_X} > x_0 > x_{\overline{F}_Y}$. Then, $\overline{F}_Y(x_0) = 1$, and since all the cumulative distribution function are 0-1 valued, F_2 $\mathcal{F}_Y$ such that $F_2(x_0) = 1$. On the other hand, $\underline{F}_X(x_0) = 0$, and since all thecumulative distribution functions are 0-1 valued, thereis some F_1 $\mathcal{F}_X$ such that $F_1(x_0) = 0$. Hence, $F_1 \leq F_2$ and therefore $\mathcal{F}_X$ $_{FSD_4}$ $\mathcal{F}_Y$.

   In this case, the preference may be non-strict.Forinstance, if $\mathcal{F}_X = \mathcal{F}_Y = \{F_1, F_2\}$ such that $x_{F_1} = 0$ and $x_{F_2} = 1$,then $x_{\underline{F}_X} = 1 > 0 = x_{\overline{F}_Y}$ but $\mathcal{F}_X \equiv_{FSD_4} \mathcal{F}_Y$.

5. If $x_{\overline{F}_X} > x_{\overline{F}_Y}$, there is some x_0 such that $x_{\overline{F}_X} > x_0 > x_{\overline{F}_Y}$. Hence, $\overline{F}_Y(x_0) = 1$. Since all the cumulative distribution functions are $0-1$ valued, $\exists F_2 \in \mathcal{F}_Y$ such that $F_2(x_0) = 1$. On the other hand, $\overline{F}_X(x_0) = 0$, whence $F_1(x_0) = 0$ for all $F_1 \in \mathcal{F}_X$. Hence, $F_1 \succ_{FSD} F_2$ for all $F_1 \in \mathcal{F}_X$. We conclude that $\mathcal{F}_X \succ_{FSD_5} \mathcal{F}_Y$ but $\mathcal{F}_Y \not\succ_{FSD_5} \mathcal{F}_X$.

   On the other hand, when $\mathcal{F}_X \succ_{FSD_5} \mathcal{F}_Y$ Proposition 4.19 implies $\overline{F}_X \succeq_{FSD} \overline{F}_Y$, and the preference must be strict because otherwise $\mathcal{F}_X$ and $\mathcal{F}_Y$ would be indifferent. Then, $x_{\overline{F}_X} > x_{\overline{F}_Y}$.

6. On the one hand, if $x_{\overline{F}_X} > x_{\overline{F}_Y}$, the result follows from the previous item and Proposition 4.15. On the other hand, when $\mathcal{F}_X \succ_{FSD_6} \mathcal{F}_Y$, Proposition 4.19 assures that $\overline{F}_X \succeq_{FSD} \overline{F}_Y$, and the preference must be strict because otherwise $\mathcal{F}_X$ and $\mathcal{F}_Y$ would be indifferent. Then, as we saw in the previous item, it holds that $x_{\overline{F}_X} > x_{\overline{F}_Y}$. ∎

Next example shows that the converse implications may not hold in general.

Example 4.57 *We begin by considering the first item. Consider the following sets of distribution functions:*

$$\mathcal{F}_X = \{F_{1,0.5-\frac{1}{n}} : n > 3\} \text{ and } \mathcal{F}_Y = \{F_{1,0.5}\}.$$

It holds that $\underline{F}_X = \underline{F}_Y = \overline{F}_Y = F_{1,0.5}$, *and then* $x_{\underline{F}_X} = x_{\underline{F}_Y} = 0.5$, *but* $\mathcal{F}_X \succ_{FSD_i} \mathcal{F}_Y$ *for* $i = 2,3,4$.

Similarly, we can consider the following sets:

$$\mathcal{F}_X = \{F_{1,0.5+\frac{1}{n}} : n > 3\} \text{ and } \mathcal{F}_Y = \{F_{1,0.5}\}.$$

It holds that $\overline{F}_X = \overline{F}_Y = F_{1,0.5}$ *and consequently* $x_{\overline{F}_X} = x_{\overline{F}_Y} = 0.5$ *but* $\mathcal{F}_X \succ_{FSD_i} \mathcal{F}_Y$ *for* $i = 5,6$.

We move next to the second item. It is enough to consider a 0-1 valued distribution function F_1 *and the sets* $\mathcal{F}_X = \mathcal{F}_Y = \{F_1\}$. *Both sets are indifferent for Definition* (FSD_i) *for* $i = 1,\ldots,6$, *but no strict inequality hold.*

Next we are going to compare the preferences between two sets of 0-1 valued distribution functions and their convex hull. Consider $S_X, S_Y \subseteq [0,1]$ and let us define the sets:

$$\mathcal{F}_{S_X} = \{F \text{ 0–1 c.d.f.} \mid x_F \in S_X\}.$$
$$\mathcal{F}_{S_Y} = \{F \text{ 0–1 c.d.f.} \mid x_F \in S_Y\}.$$

Since we are working with σ-additive cumulative distribution functions, $\mathcal{F}_{S_X}$ and $\mathcal{F}_{S_Y}$ are related to the degenerate probability measures on elements of S_X, S_Y, respectively.

We shall also consider theirconvex hulls $\mathcal{F}_X := conv(\mathcal{F}_{S_X})$, $\mathcal{F}_Y := conv(\mathcal{F}_{S_Y})$. These are the sets of cumulative distribution functions with finite supp orts that are included in S_X and S_Y, resp ectively.

Now, given any set $\mathcal{F}$ ofcumulativedistributionfunctionsand itsconvexhull $\mathcal{F}_c$, the p-b oxes $(\underline{F}, \overline{F})$ and $(\underline{F}_c, \overline{F}_c)$ asso ciated with $\mathcal{F}, \mathcal{F}_c$, coincide:

$$\underline{F} = \underline{F}_c \quad \overline{F} = \overline{F}_c. \tag{4.6}$$

Thus, $\mathcal{F}_X$ and $\mathcal{F}_{S_X}$ determine the same p-b ox, and the same applies to $\mathcal{F}_Y$ and $\mathcal{F}_{S_Y}$. We begin with an immediate lemma, whose pro of is trivial and therefore omitted.

Lemma 4.58 *Consider* $S \subseteq [0, 1]$ *and* $\mathcal{F}_S = \{F$ *0-1 c.d.f.* $\mid x_F \quad S\}$. *Let* $\underline{x} = \inf S$ *and* $\overline{x} = \sup S$ *and let* $\underline{F}, \overline{F}$ *be the lower and upper distribu tion functions associated with* $\mathcal{F}$. *Then*

$$\underline{F} = I_{[\overline{x}, 1]} \text{ and } \overline{F} = \begin{cases} I_{[\underline{x}, 1]} & \text{if } \underline{x} \quad S, \\ I_{(\underline{x}, 1]} & \text{otherwise.} \end{cases}$$

Moreover, if $\overline{x} \quad S$, *then* $\underline{F} \quad \mathcal{F}$, *and if* $\underline{x} \quad S$, *then* $\overline{F} \quad \mathcal{F}$.

Note that when $\overline{F} = I_{(\underline{x}, 1]}$, this is a finite, but not cumulative, distribution function, and as a con sequence it cannot b elong to $\mathcal{F}_S$.

Prop osition 4.59 *Let* S_X *and* S_Y *be twosubsets of* $[0, 1]$ *Then:*

1. $\mathcal{F}_X \quad_{FSD_1} \mathcal{F}_Y \quad \mathcal{F}_{S_X} \quad_{FSD_1} \mathcal{F}_{S_Y} \quad \inf S_X \geq \sup S_Y$.

If in addition both $\inf S_X$ *and* $\sup S_X$ *belong to* S_X, *and also* $\inf S_Y$ *and* $\sup S_Y$ *belong to* S_Y, *then also:*

2. $\mathcal{F}_X \quad_{FSD_2} \mathcal{F}_Y \quad \mathcal{F}_{S_X} \quad_{FSD_2} \mathcal{F}_{S_Y} \quad \max S_X \geq \max S_Y$. *Moreover,*
   $\max S_X > \max S_Y \quad \mathcal{F}_{S_X} \quad_{FSD_2} \mathcal{F}_{S_Y}$ *and*
   $\max S_X = \max S_Y \quad \mathcal{F}_{S_X} \equiv_{FSD_2} \mathcal{F}_{S_Y}$.
3. $\mathcal{F}_X \quad_{FSD_3} \mathcal{F}_Y \quad \mathcal{F}_{S_X} \quad_{FSD_3} \mathcal{F}_{S_Y} \quad \max S_X \geq \max S_Y$. *Moreover,*
   $\max S_X > \max S_Y \quad \mathcal{F}_{S_X} \quad_{FSD_3} \mathcal{F}_{S_Y}$ *and*
   $\max S_X = \max S_Y \quad \mathcal{F}_{S_X} \equiv_{FSD_3} \mathcal{F}_{S_Y}$.
4. $\mathcal{F}_X \quad_{FSD_4} \mathcal{F}_Y \quad \mathcal{F}_{S_X} \quad_{FSD_4} \mathcal{F}_{S_Y} \quad \max S_X \geq \min S_Y$. *Moreover,*
   $\max S_X > \min S_Y \quad \mathcal{F}_{S_X} \quad_{FSD_4} \mathcal{F}_{S_Y}$ *and*
   $\max S_X = \min S_Y \quad \mathcal{F}_{S_X} \equiv_{FSD_4} \mathcal{F}_{S_Y}$.
5. $\mathcal{F}_X \quad_{FSD_5} \mathcal{F}_Y \quad \mathcal{F}_{S_X} \quad_{FSD_5} \mathcal{F}_{S_Y} \quad \min S_X \geq \min S_Y$. *Moreover,*
   $\min S_X > \min S_Y \quad \mathcal{F}_{S_X} \quad_{FSD_5} \mathcal{F}_{S_Y}$ *and*
   $\min S_X = \min S_Y \quad \mathcal{F}_{S_X} \equiv_{FSD_5} \mathcal{F}_{S_Y}$.

6. $\mathcal{F}_X \quad_{FSD_6} \mathcal{F}_Y \quad \mathcal{F}_{S_X} \quad_{FSD_6} \mathcal{F}_{S_Y} \quad \min S_X \geq \min S_Y$. *Moreover,* $\min S_X > \min S_Y \quad \mathcal{F}_{S_X} \quad_{FSD_6} \mathcal{F}_{S_Y}$ *and* $\min S_X = \min S_Y \quad \mathcal{F}_{S_X} \equiv_{FSD_6} \mathcal{F}_{S_Y}$.

Proof The first statement follows from Proposition 4.19 and Equation (4.6), taking also into account that, from Lemma 4.58, $\overline{F}_X \leq \underline{F}_Y$ if and only if $\inf S_X \geq \sup S_Y$.

To prove the other statements, note first of all that if the infima and suprema of S_X and S_Y are included in the set, it follows from Lemma 4.58 that $\underline{F}_X, \overline{F}_X \quad \mathcal{F}_{S_X}$ and $\underline{F}_Y, \overline{F}_Y \quad \mathcal{F}_{S_Y}$, and applying Corollary 4.22 together with Equation (4.6) we deduce that

$$\mathcal{F}_X \quad_{FSD_i} \mathcal{F}_Y \quad \mathcal{F}_{S_X} \quad_{FSD_i} \mathcal{F}_{S_Y} \quad i = 2, \ldots, 6.$$

On the other hand, it follows from Lemma 4.58 that in those cases

$$\underline{F}_X = I_{[\max S_X, 1]}, \ \underline{F}_Y = I_{[\max S_Y, 1]}, \ \overline{F}_X = I_{[\min S_X, 1]}, \ \overline{F}_Y = I_{[\min S_Y, 1]}.$$

The second and third equivalences in each statement follow then from Corollary 4.22. ∎

As a consequence of this result, we obtain the following corollary.

Corollary 4.60 *If S_X and S_Y are closed subsets of $[0, 1]$ then:*

1. $\mathcal{F}_{S_X} \quad_{FSD_1} \mathcal{F}_{S_Y} \quad \min S_X \geq \max S_Y$.
2. $\mathcal{F}_{S_X} \quad_{FSD_2} \mathcal{F}_{S_Y} \quad \max S_X \geq \max S_Y$.
3. $\mathcal{F}_{S_X} \quad_{FSD_3} \mathcal{F}_{S_Y} \quad \max S_X \geq \max S_Y$.
4. $\mathcal{F}_{S_X} \quad_{FSD_4} \mathcal{F}_{S_Y} \quad \max S_X \geq \min S_Y$.
5. $\mathcal{F}_{S_X} \quad_{FSD_5} \mathcal{F}_{S_Y} \quad \min S_X \geq \min S_Y$.
6. $\mathcal{F}_{S_X} \quad_{FSD_6} \mathcal{F}_{S_X} \quad \min S_X \geq \min S_Y$.

Hence, in that case (FSD_2) *is equivalent to* (FSD_3) *and* (FSD_5) *is equivalent to* (FSD_6).

It is easy to see that Proposition 4.59 and Corollary 4.60 also hold when we consider $\mathcal{F}_X$ and $\mathcal{F}_Y$ given by

$$\mathcal{F}_X = \{F \text{ c.d.f.} \mid P_F(S_X) = 1\} \text{ and } \mathcal{F}_Y = \{F \text{ c.d.f.} \mid P_F(S_Y) = 1\}.$$

4.1.2 Imprecise statisticalpreference

In Section 4.1.1 we considered the particular case in which the binary relation is sto chastic dominance. Now we fo cus on the case where the binary relation is that of s tati stical preference, givenin Definition 2.16. Hence, we shall assume that the utility space Ω is an ordered set, which ne ed not b e numeric al.

Remark 4.61 *Analogously to the case of stochastic dominance, we shal l denote by* $_{SP_i}$, $i = 1, \ldots, 6$ *the conditions obtainedby using statistical preference asthe binary relation in Definition 4.1. We shal l also say that* X *is* (SP_i) *preferred or* (SP_i) *statistical ly preferred to* Y *when* $X \ _{SP_i} Y$. *Furthermore, thenotation* $X \ _{SP_{i,j}} Y$ *means that* $X \ _{SP_i} Y$ *and* $X \ _{SP_j} Y$. *Notethat inSection4.1.1weusedinterchangeablythenotation* $X \ _{FSD_i} Y$ *and* $F_X \ _{FSD_i} F_Y$, *since stochastic dominance isbased on the directcomparison of the cumulative distribution functions. Now, we shall only employ the notation* $X \ _{SP_i} Y$, *because statisticalpreferenceis based on the joint distribution of the random variables, and the marginal distributions do not keep all the information about it.*

When the binary relation is sto chastic dominance, we saw in Prop osition 4.15 that there are some general relationships between its strict extensions. Inthe case of statistic al preference,the relationships showed in Prop osition 4.15 do not hold in general, as we cansee fromthe following example:

Example 4.62 *Considerthe universe* $\Omega = \{\omega_1, \omega_2, \omega_3\}$ *and let* P *bethediscrete uniform distribution on* Ω. *Considerthe setsof random variables* $\mathcal{X} = \{X_1, X_2, X_3\}$ *and* $\mathcal{Y} = \{X_2, X_4\}$, *where the randomvariables are defined by:*

	ω_1	ω_2	ω_3
X_1	0	2	4
X_2	4	0	2
X_3	2	4	0
X_4	3	2	1

For these sets, since $X_1 \ _{SP} X_2$ *and* $X_1 \equiv_{SP} X_4$, *then* $\mathcal{X} \ _{SP_2} \mathcal{Y}$. *Moreover, since* $X_2 \ _{SP} X_1$ *and* $X_4 \ _{SP} X_2$, *we have that* $\mathcal{Y} \ _{SP_2} \mathcal{X}$, *hence* $\mathcal{X} \ _{SP_2} \mathcal{Y}$.

However, $\mathcal{X} \ _{SP_3} \mathcal{Y}$: *since* $X_1 \equiv_{SP} X_4$, $X_2 \equiv_{SP} X_2$ *and* $X_4 \ _{SP} X_3$, *it holds that* $\mathcal{Y} \ _{SP_3} \mathcal{X}$. *Hence,* $\mathcal{X} \equiv_{SP_3} \mathcal{Y}$.

With asimilar example it could be proved that $\mathcal{X} \ _{SP_5} \mathcal{Y}$ *and* $\mathcal{X} \equiv_{SP_6} \mathcal{Y}$ *are compatible st at ement s.*

Note that $_{SP}$ is reflexive and comple te,but it is ne ither antisymmetric nor transitive. Hence, Prop osition 4.6 do es not apply in this case; indeed, we can use statistical prefer- ence to show that Prop osition 4.6 cannot be extended to non transitive relationships.

Example 4.63 *Considerthe random variables A, B, C from Example3.83 such that $A \;_{SP}\; B \;_{SP}\; C \;_{SP}\; A$, and let $X = \{A, B\}, Y = \{A, C\}$. Then since $A \;_{SP}\; A$ and $B \;_{SP}\; C$, we deducethat $X \;_{SP_3}\; Y$; since $A \;_{SP}\; B$ and $C \;_{SP}\; A$, we see that $X \;_{SP_2}\; Y$; however, X has a maximumelement,because $A \;_{SP}\; B$.*

On the other hand, since statistical preference complies with Pareto dominance we deduce from Prop osition 4.7 that the different conditions can be reduced to the comparison of the maximum andminimum elements of X, Y, when these maximum and minimum elements exist. Finally, we deduce from Prop ositions 4.9 and 4.11 that conditions $_{SP_3}$, $_{SP_4}$, $_{SP_6}$ induce a reflexive and comple te relationship.

We can also use statistical preference to show that Prop osition 4.11 cannot be extended to the relations $_1$, $_2$ nor $_5$: take thesets $X = Y = \{A, B, C\}$, where the variables A, B, C satisfy $A \;_{SP}\; B \;_{SP}\; C \;_{SP}\; A$ as in Example 3.83; then the set X has neither a maximum nor a minimu m element, whence it is incomparable with itself with resp ect to $_{SP_2}$ and $_{SP_5}$. Applying Prop osition 4.3, we deduce that X, Y are also incomparable with resp ect to $_{SP_1}$.

Weshowedin Theorem 4.23 thatthe generalisationsof sto chastic dominance towards sets of variables are related to lower and upp er exp ectations. Next, weestablisha similar resultforthe generalisationsof statisticalpreference. RecallthatinTheorem 3.40 we proved that:

$$\sup \mathrm{Me}(X - Y) > 0 \qquad X \;_{SP}\; Y \qquad \sup \mathrm{Me}(X - Y) \geq 0. \tag{4.7}$$

Taking into thisresult, we shall establish a generalisation in terms of lower and upp er medians, andfor thisweshall requireour utility space Ω to be the reals. Let usconsider two sets ofalternatives X, Y with valueson Ω, and let us intro duce the following notation:

$$\begin{aligned} \mathrm{Me}(X - Y) &= \{\mathrm{Me}(X - Y) : X \quad X, Y \quad Y\}. \\ \underline{\mathrm{Me}}(X - Y) &= \inf \mathrm{Me}(X - Y). \\ \mathrm{Me}(X - Y) &= \sup \mathrm{Me}(X - Y), \end{aligned}$$

where we recall that the median of a random variable with resp ect to a probability measureis given by Equation (3.14).

Prop osition 4.64 *Let X, Y betwosets of random variables defined on a probability space $(\Omega, \mathcal{A}, P)$ and taking values on $\mathbb{R}$.*

1. $\underline{\mathrm{Me}}(X - Y) > 0 \quad X \;_{SP_1}\; Y \quad \overline{\mathrm{Me}}(X - Y) \geq 0.$
2. $X \quad X$ *such that* $\underline{\mathrm{Me}}(\{X\} - Y) > 0 \quad X \;_{SP_2}\; Y \quad X \quad X$ *such that* $\overline{\mathrm{Me}}(\{X\} - Y) \geq 0.$
3. $\overline{\mathrm{Me}}(X - \{Y\}) > 0 \quad Y \quad Y \quad X \;_{SP_3}\; Y \quad \overline{\mathrm{Me}}(X - \{Y\}) \geq 0 \quad Y \quad Y.$

4. $\overline{\mathrm{Me}}(X - Y) > 0 \quad X \quad_{SP_4} Y \quad \overline{\mathrm{Me}}(X - Y) \geq 0.$
5. $Y \quad Y$ *such that* $\underline{\mathrm{Me}}(X - \{ Y\}) > 0 \quad X \quad_{SP_5} Y \quad Y \quad Y$ *such that* $\overline{\mathrm{Me}}(X - \{Y\}) \geq 0.$
6. $\overline{\mathrm{Me}}(\{X\} - Y) > 0 \quad X \quad X \quad X \quad_{SP_6} Y \quad \overline{\mathrm{Me}}(\{X\} - Y) \geq 0 \quad X \quad X.$

Pro of Recalloncemore thatfromEquation(4.7) giventworandomvariables X, Y,

$$\overline{\mathrm{Me}}(X - Y) > 0 \quad X \quad_{SP} Y \quad \overline{\mathrm{Me}}(X - Y) \geq 0.$$

SP1 : If $\underline{\mathrm{Me}}(X - Y) > 0$, in particu lar $\overline{\mathrm{Me}}(X - Y) > 0$, and th en $\mathrm{Me}(X - Y) > 0$ for every $X \quad X$ and $Y \quad Y$. Applying Equation(4.7), $X \quad_{SP} Y$ for every $X \quad X$ and $Y \quad Y$, and consequently $X \quad_{SP_1} Y$. Moreover,

$$X \quad_{SP_1} Y \quad X \quad_{SP} Y \text{ for every } X \quad X, Y \quad Y$$
$$\sup \mathrm{Me}(X - Y) \geq 0 \text{ for every } X \quad X, Y \quad Y \quad \overline{\mathrm{Me}}(X - Y) \geq 0.$$

SP2 : If there is some $X \quad X$ such that $\underline{\mathrm{Me}}(\{X\} - Y) > 0$, then $\mathrm{Me}(X - Y) > 0$ for every $Y \quad Y$. Applying Equation(4.7), wededuce that $X \quad_{SP} Y$ for every $Y \quad Y$, and therefore $X \quad_{SP_2} Y$.

On the other hand,

$$X \quad_{SP_2} Y \quad \text{there is some } X \quad X \text{ such that } X \quad_{SP} Y \text{ for every } Y \quad Y$$
$$\sup \mathrm{Me}(X - Y) \geq 0 \text{ for every } Y \quad Y \quad \mathrm{Me}(\{X\} - Y) \geq 0.$$

SP3 : Consider $Y \quad Y$. If $\overline{\mathrm{Me}}(X - \{ Y\}) > 0$, then there is some $X \quad X$ such that $\mathrm{Me}(X - Y) > 0$. Hence, for every $Y \quad Y$ there is $X \quad X$ such that $X \quad_{SP} Y$, and consequently $X \quad_{SP_3} Y$. Moreover,

$$X \quad_{SP_3} Y \quad \text{for every } Y \quad Y \text{ there is } X \quad X \text{ such that } X \quad_{SP} Y$$
$$\text{for every } Y \quad Y \text{ there is } X \quad X \text{ such that } \sup \mathrm{Me}(X - Y) \geq 0$$
$$\text{for every } Y \quad Y \text{ itholds that } \mathrm{Me}(X - \{ Y\}) \geq 0.$$

SP4 : If $\overline{\mathrm{Me}}(X - Y) > 0$, there are $X \quad X$ and $Y \quad Y$ such that $\mathrm{Me}(X - Y) > 0$, and consequently $X \quad_{SP} Y$. Thus, $X \quad_{SP_4} Y$. On theotherhand,

$$X \quad_{SP_4} Y \quad \text{there are } X \quad X, Y \quad Y \text{ such that } X \quad_{SP} Y$$
$$\text{there are } X \quad X, Y \quad Y \text{ such that } \sup \mathrm{Me}(X - Y) \geq 0 \quad \overline{\mathrm{Me}}(X - Y) \geq 0.$$

SP5 : Assume that the re exists some $Y \quad Y$ such that $\underline{\mathrm{Me}}(X - \{ Y\}) > 0$. Then $\mathrm{Me}(X - Y) > 0$ for every $X \quad X$, andapplying (4.7) weconcludethatthereis $Y \quad Y$

such that X $_{SP}$ Y for every X X, and consequently X $_{SP_5}$ Y. On the other hand,

X $_{SP_5}$ Y there is Y Y such that X $_{SP}$ Y for every Y Y
there is Y Y such that $\underline{\sup}$ $\mathrm{Me}(X - Y) \geq 0$ for every X X
there is Y Y such that $\mathrm{Me}(X - \{Y\}) \geq 0$.

SP6 : Finally, if $\overline{\mathrm{Me}}(\{X\} - Y) > 0$ for every X X, then for every X X there is some Y Y such that $\mathrm{Me}(X - Y) > 0$, whence (4.7) implies that X $_{SP}$ Y. We conclude that X $_{SP_6}$ Y. Moreover,

X $_{SP_6}$ Y for every X X there is Y Y such that X $_{SP}$ Y
for every X X there is Y Y such that $\sup \mathrm{Me}(X - Y) \geq 0$
for every X X, $\mathrm{Me}(\{X\} - Y) \geq 0$.

Taking into account the properties of the median, we conclude from this result that statistical preference may be seen as a more robust alternative to stochastic dominance or expected utility in the presence of outliers.

As we made in Section 4.1.1 with imprecise stochastic dominance, now we shall investigate some of the properties of the imprecise statistical preference.

Increasing imprecision

We first study the behavior of conditions $_{SP_i}$, $i = 1, \ldots, 6$, when we enlarge the sets X, Y of alternatives we want to compare. This may correspond to an increase in the imprecision of our models. Not surprisingly, if the more restrictive condition $_{SP_1}$ is satisfied on the large sets, then it is automatically satisfied on the smaller ones; while for the least restrictive one $_{SP_4}$ we have the opposite implication.

Proposition 4.65 *Let X, Y, X and Y be four sets of random variables satisfying X X and Y Y. Then*

$$X \quad _{SP_1} Y \quad X \quad _{SP_1} Y \text{ and } X \quad _{SP_4} Y \quad X \quad _{SP_4} Y.$$

Proof It is clear that X $_{SP_1}$ Y X $_{SP_1}$ Y, since if X $_{SP}$ Y for every X X and Y Y, the inequality holds in particular for every X X and Y Y.

On the other hand, X $_{SP_4}$ Y implies the existence of X X and Y Y satisfying X $_{SP}$ Y, and then the inclusions X X and Y Y imply that X $_{SP_4}$ Y. ∎

Similar implications cannot be established for $_{SP_i}$, for $i = 2, 3, 5, 6$, as the following example shows:

Example 4.66 *Consider theuniverse $\Omega = \{\omega\}$ and let δ_x denotetherandom variable satisfying $\delta_x(\omega) = x$.*

Let us prove that $X \ {}_{SP_i}\ Y$ and $Y \ {}_{SP_i}\ X$ is possiblefor $i = 2, 3, 5, 6$:

- *Consider $X = \{\delta_0\}, X = \{\delta_0, \delta_2\}$ and $Y = Y = \{\delta_1\}$. It holdsthat $Y \ {}_{SP_i}\ X$ for $i = 1, \ldots, 6$ while $X \ {}_{SP_i}\ Y$ for $i = 2, 3$, since $\delta_2 \ {}_{SP}\ \delta_1$.*
- *Now, given $X = \{\delta_2\}, X = \{\delta_0, \delta_2\}$ and $Y = Y = \{\delta_1\}$, itholds that $X \ {}_{SP_i}\ Y$ for $i = 1, \ldots, 6$ while $Y \ {}_{SP_i}\ X$ for $i = 5, 6$, since $\delta_1 \ {}_{SP}\ \delta_0$.*

Note that these examples also show that the implications of the previous proposition are not equivalences in general.

One particular case when wemay enlargeour sets of alternatives is when we consider convex combinations (note that for this we shall again to assume that the utility space Ω is equalto $\mathbb{R}$). This may be of interest for instance if we want to compare random sets by means of their measurable selections, as weshall do in Section 4.2.1, and wemove from a purely atomic toa non-atomic initial probability space. We shall consider two possibilities, for a given set of alternatives D: its *convex hull*

$$Conv(D) = \left\{ U = \sum_{i=1}^{n} \lambda_i X_i : \lambda_i > 0, X_i \ D \ i, \sum_{i=1}^{n} \lambda_i = 1 \right\},$$

and also the se t of alternatives whose utilities b elong to the range of utilities determined by A:

$$Conv(D) = \{U \text{ r.v.} \mid U(\omega) \ Conv(\{U(\omega) : U \ D\})\}; \tag{4.8}$$

note that $D \ \ Conv(D) \ \ Conv(D)$. Then Prop osition 4.65 allows to immediately deduce the following:

Corollary 4.67 *Consider twosets of alternatives X, Y.*

(a) $Conv(X) \ {}_{SP_1}\ Conv(Y) \quad Conv(X) \ {}_{SP_1}\ Conv(Y) \quad X \ {}_{SP_1}\ Y.$

(b) $Conv(X) \ {}_{SP_4}\ Conv(Y) \quad X \ {}_{SP_4}\ Y \quad Conv(X) \ {}_{SP_4}\ Conv(Y).$

To see that we cannot establish similar implications with resp ect to ${}_{SP_i}$, $i = 2, 3, 5, 6$, take the following example:

Example 4.68 *Consider* $\Omega = \{\omega_1, \omega_2, \omega_3\}$ *with* $P(\{\omega_i\}) = \frac{1}{3}$ *for every* $i = 1,2,3$. *Let us consider the sets of variables* $\mathcal{X} = \{X_1, X_2\}$ *and* $\mathcal{Y} = \{Y\}$ *given by:*

	ω_1	ω_2	ω_3
X_1	0	3	0
X_2	3	0	0
Y	1	1	1

Then since $Q(X_1, Y) = Q(X_2, Y) = \frac{1}{3}$, *it follows that* $\mathcal{Y} \ {}_{SP_i} \ \mathcal{X}$ *for* $i = 1, \ldots, 6$. *However,* $Conv(\mathcal{X}) \ {}_{SP_i} \ Conv(\mathcal{Y})$, *for* $i = 2,3$, $Conv(\mathcal{X}) \equiv_{SP_4} Conv(\mathcal{Y})$ *and they are incomparable with respect to* ${}_{SP_1}$.

On the other hand, if we consider instead the sets $\mathcal{X} = \{X_1, X_2\}$ *and* $\mathcal{Y} = \{Y\}$, *where*

	ω_1	ω_2	ω_3
X_1	0	3	3
X_2	3	0	3
Y	2	2	2

it holds that $\mathcal{X} \ {}_{SP_i} \ \mathcal{Y}$ *for* $i = 1, \ldots, 6$. *However,* $Conv(\mathcal{Y}) \ {}_{SP_i} \ Conv(\mathcal{X})$, *for* $i = 5,6$.

The same sets of variables show that there is no additional implication if we consider the convex hulls determined by Equation (4.8) *instead.*

Connection with aggregation functions

Since the binary relation associated with statistical preference is complete, we deduce from Proposition 4.11 that the relations ${}_{SP_3}$, ${}_{SP_4}$, ${}_{SP_6}$ also induce a complete relation. Such relations are interesting because they mean that we can always express a preference between two sets of alternatives $\mathcal{X}, \mathcal{Y}$. One way of deriving a complete relation when we make multiple comparisons is to establish a degree of preference for every pairwise comparison, and to aggregate these degrees of preference into a joint one. This is possible by means of an aggregation function.

Let $\mathcal{X} = \{X_1, \ldots, X_n\}$ and $\mathcal{Y} = \{Y_1, \ldots, Y_m\}$ be two finite sets of random variables taking values on an ordered utility space Ω, and let us compute the statistical preference $Q(X_i, Y_j)$ for every pair of variables $X_i \in \mathcal{X}$, $Y_j \in \mathcal{Y}$ by means of Equation (2.7). The set of all these preferences is an instance of *profile of preference* [80], and can be represented by means of the matrix

$$\mathrm{Q}^{\mathcal{X},\mathcal{Y}} := \begin{pmatrix} Q(X_1,Y_1) & Q(X_1,Y_2) & \cdots & Q(X_1,Y_m) \\ \vdots & \vdots & \ddots & \vdots \\ Q(X_n,Y_1) & Q(X_n,Y_2) & \cdots & Q(X_n,Y_m) \end{pmatrix} \tag{4.9}$$

Note that the profile of preferences of Y over X, $Q^{Y,X}$, corresponds to one minus the transposed matrix of $Q^{X,Y}$, i.e., $1 - Q^t_{X,Y}$. We shall show that conditions $SP_1, \ldots, SP_6$ can be expressed by means of an aggregation function over the profile of preference:

Definition 4.69 ([31, 80]) *An* aggregation function *is a mapping defined by*

$$G: \bigcup_{s \in \mathbb{N}} [0, 1]^s \to [0, 1],$$

that it componentwise increasing and satisfies the boundary conditions $G(0, \ldots, 0) = 0$ *and* $G(1, \ldots, 1) = 1$.

The matrix $Q^{X,Y}$ representing the profile of preferences between X and Y can be equivalently represented by means of a vector on $[0, 1]^{nm}$ using the lexicographic order:

$$z_{X,Y} = (Q(X_1, Y_1), Q(X_1, Y_2), \ldots, Q(X_1, Y_m), Q(X_2, Y_1), \ldots, Q(X_n, Y_m)).$$

Taking this into account, given an aggregation function $G: \bigcup_{s \in \mathbb{N}} [0, 1]^s \to [0, 1]$ we shall denote by $G(Q^{X,Y})$ the image of the vector $z_{X,Y}$ by means of this aggregation function.

Definition 4.70 *Given two finite sets of random variables X and Y, $X = \{X_1, \ldots, X_n\}$ and $Y = \{Y_1, \ldots, Y_m\}$, and an aggregation function G, we say that X is G*-statistically preferred to *Y, and denote it by $X \succeq_{SP_G} Y$, if*

$$G(Q^{X,Y}) := G(z_{X,Y}) \geq \frac{1}{2}. \tag{4.10}$$

We refer to [31] for a review of aggregation functions. Some important properties are the following:

Definition 4.71 ([31]) *An aggregation function $G: \bigcup_{s \in \mathbb{N}} [0, 1]^s \to [0, 1]$ is called:*

- Symmetric *if it is invariant under permutations.*
- Monotone *if $G(r_1, \ldots, r_s) \geq G(r'_1, \ldots, r'_s)$ whenever $r_i \geq r'_i$ for every $i = 1, \ldots, s$.*
- Idempotent *if $G(r, \ldots, r) = r$.*

We shall call an aggregation function $G: \bigcup_{s \in \mathbb{N}} [0, 1]^s \to [0, 1]$ *self-dual* if

$$G(r_1, \ldots, r_s) = 1 - G(1 - r_1, \ldots, 1 - r_s)$$

for every $(r_1, \ldots, r_s) \in [0, 1]^s$ and for every $s \in \mathbb{N}$.

All these properties are interesting when aggregating the profile of preferences into a joint one: symmetry implies that all the elements in the profile are given the same weight; idemp otency means that if all the preference degrees equal t, the final preference degree should also equal t; monotonicity assures that if we increase all the values in the profile of preferences , the final valu e should also increase, and self-dualitypreserves the idea behind the notion of probabilistic relation in De finition 2.7, since fora self-dual aggregation function G, $G(Q^t_{X,Y}) + G(Q_{Y,X}) = 1$. If in addition G is symmetric, we obtain that $G(Q_{X,Y}) + G(Q_{Y,X}) = 1$.

This last prop erty means that, when G is a self-dual and symmetric aggregation function, Equation (4. 10) is equivalent to $G(Q_{X,Y}) \geq G(Q_{Y,X})$.

The relations $_{SP_i}$, for $i = 1, \ldots, 6$, can all expressed by means of an aggregation function, as we summarise in the following prop osition. Its pro of is immediate and therefore omitted.

Prop osition 4.72 *Let $X = \{X_1, \ldots, X_n\}$, $Y = \{Y_1, \ldots, Y_m\}$ be two finite sets of random variables taking values onan ordered space Ω . Thenfor any $i = 1, \ldots, 6$ X $_{SP_i}$ Y if and only if it is G_i-statistical ly preferred to Y, where the aggregationfunctions G_i are given by:*

$$G_1(Q_{X,Y}) := \min_{i,j} Q(X_i, Y_j).$$

$$G_2(Q_{X,Y}) := \max_{i=1,\ldots,n} \min_{j=1,\ldots,m} Q(X_i, Y_j).$$

$$G_3(Q_{X,Y}) := \min_{j=1,\ldots,m} \max_{i=1,\ldots,n} Q(X_i, Y_j).$$

$$G_4(Q_{X,Y}) := \max_{i,j} Q(X_i, Y_j).$$

$$G_5(Q_{X,Y}) := \max_{j=1,\ldots,m} \min_{i=1,\ldots,n} Q(X_i, Y_j).$$

$$G_6(Q_{X,Y}) := \min_{i=1,\ldots,n} \max_{j=1,\ldots,m} Q(X_i, Y_j).$$

It is not difficult to see that all the aggregationfunctions G_i ab ove are monotonic and comply with the boundary conditions $G_i(0, \ldots, 0) = 0$ and $G_i(1, \ldots, 1) = 1$. On the other hand, only G_1 and G_4 are symmetric, and none of the m is self-dual.

We can also use these aggregation functions to deduce the relationships between the different conditions established in Prop osition 4.3 in the case of statistical preference: it suffices to takeinto account that $G_1 \leq G_2 \leq G_3 \leq G_4$ and $G_1 \leq G_5 \leq G_6 \leq G_4$.

Remark 4.73 *Proposition4.72 helpsto verify eachof theconditions* $_{SP_i}$, $i = 1, \ldots, 6$ *by looking at the profileof preferences $Q_{X,Y}$ given byEquation* (4.9)*:*

- *X $_{SP_1}$ Y if and only if all elements in the mat rix are greater than or equalto $\frac{1}{2}$.*

- $X \succeq_{SP_2} Y$ *if and only if there is a row whose elements are all greater than or equal to* $\frac{1}{2}$.
- $X \succeq_{SP_3} Y$ *if and only if in each column there is at least one element greater than or equal to* $\frac{1}{2}$.
- $X \succeq_{SP_4} Y$ *if and only if there is an element greater than or equal to* $\frac{1}{2}$.
- $X \succeq_{SP_5} Y$ *if and only if there is a column whose elements are all greater than or equal to* $\frac{1}{2}$.
- $X \succeq_{SP_6} Y$ *if and only if in each row there is at least one element greater than or equal to* $\frac{1}{2}$.

See the comments after Proposition 4.3 for a related idea.

The above remarks suggest that other preference relationships may be defined by means of other aggregation functions G, and this would allow us to take all the elements of the profile of preferences into account, instead of focusing on the best or worst scenarios only. Next, we explore briefly one of these possibilities: the arithmetic mean G_{mean}, given by

$$\begin{aligned} G_{\text{mean}} : \bigcup_{s \in \mathbb{N}} [0,1]^s &\to [0,1] \\ (r_1, \ldots, r_s) &\to \frac{r_1 + \cdots + r_s}{s}. \end{aligned}$$

This is a symmetric, monotone, idempotent and self-dual aggregation function. For clarity, when X is G_{mean}-statistically preferred to Y we shall denote it $X \succeq_{SP_{\text{mean}}} Y$. The connection between $\succeq_{SP_{\text{mean}}}$ and $\succeq_{SP_i}$, $i = 1, \ldots, 6$ is a consequence of the following result:

Proposition 4.74 *Given two finite sets of random variables* $\mathcal{X}$ *and* $\mathcal{Y}$, $\mathcal{X} = \{X_1, \ldots, X_n\}$ *and* $\mathcal{Y} = \{Y_1, \ldots, Y_m\}$, *and a monotone and idempotent aggregation function* G,

$$\mathcal{X} \succeq_{SP_1} \mathcal{Y} \Rightarrow \mathcal{X} \succeq_{SP_G} \mathcal{Y} \Rightarrow \mathcal{X} \succeq_{SP_4} \mathcal{Y}.$$

Proof On the one hand, assume that $\mathcal{X} \succeq_{SP_1} \mathcal{Y}$. Then, $Q(X, Y) \geq \frac{1}{2}$ for every $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$. Since G is monotone and idempotent, $G(\mathcal{Q}^{\mathcal{X},\mathcal{Y}}) \geq G\left(\frac{1}{2}, \ldots, \frac{1}{2}\right) = \frac{1}{2}$, and consequently $\mathcal{X} \succeq_{SP_G} \mathcal{Y}$.

On the other hand, assume ex-absurdo that $G(\mathcal{Q}^{\mathcal{X},\mathcal{Y}}) \geq \frac{1}{2}$ and that $\mathcal{X} \not\succeq_{SP_4} \mathcal{Y}$, so that $Q(X, Y) < \frac{1}{2}$ for every $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$. Then $G(\mathcal{Q}^{\mathcal{X},\mathcal{Y}}) \leq \max_{i,j} Q(X_i, Y_j) < \frac{1}{2}$, a contradiction. Hence, $\mathcal{X} \succeq_{SP_4} \mathcal{Y}$. ∎

In particular, we see that $\succeq_{SP_{\text{mean}}}$ is an intermediate notion between $\succeq_{SP_1}$ and $\succeq_{SP_4}$. To see that it is not related to $\succeq_{SP_i}$ for $i = 2, 3, 5, 6$, consider the following example:

Example 4.75 *Consider* $\Omega = \{\omega_1, \omega_2\}$ ($P(\{\omega\}) = 1/2$), *and the sets of random variables* $\mathcal{X} = \{X_1, X_2, X_3\}$ *and* $\mathcal{Y} = \{Y\}$ *defined by:*

	ω_1	ω_2
X_1	0	2
X_2	0	0
X_3	2	2
Y	1	1

Then,

$$Q^{\mathcal{X},\mathcal{Y}} := \begin{pmatrix} \frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \text{ and } Q^{\mathcal{Y},\mathcal{X}} := \begin{pmatrix} \frac{1}{2} & 1 & 0 \end{pmatrix}$$

whence Remark 4.73 implies that $\mathcal{X}$ $_{SP_i}$ $\mathcal{Y}$, *for* $i = 2,3$, *and* $\mathcal{Y}$ $_{SP_i}$ $\mathcal{X}$, *for* $i = 5,6$. *On the other hand,*

$$\frac{Q(X_1, Y) + Q(X_2, Y) + Q(X_3, Y)}{3} = \frac{1}{2},$$

and consequently $\mathcal{X} \equiv_{SP_{mean}} \mathcal{Y}$. *Hence,* $\mathcal{X}$ $_{SP_{mean}}$ $\mathcal{Y}$ $\mathcal{X}$ $_{SP_i}$ $\mathcal{Y}$ *for* $i = 5,6$, *and* $\mathcal{Y}$ $_{SP_{mean}}$ $\mathcal{X}$ $\mathcal{Y}$ $_{SP_i}$ $\mathcal{X}$ *for* $i = 2,3$. *By comparing* $\mathcal{Z}_1 = \{X_2, Y\}$ *and* $\mathcal{Z}_2 = \{X_3, Y\}$ *with* $\mathcal{X}$, *we can see that:* $\mathcal{Z}_1 \equiv_{SP_{5,6}} \mathcal{X}$ $_{SP_{mean}}$ $\mathcal{Z}_1$ *and* $\mathcal{Z}_2 \equiv_{SP_{2,3}} \mathcal{X}$ $_{SP_{mean}}$ $\mathcal{Z}_2$. *Then, there are not general relationships between* $_{SP_{mean}}$ *and* $_{SP_i}$ *for* $i = 2, 3, 5, 6$.

4.2 Modelling imprecision in decision making problems

In this section, we shall show how the above results can be applied in two different scenarios where imprecision enters a decision problem: the case where we have imprecise information about the utilities of the different alternatives, and that where we have imprecise beliefs about the states of nature.

4.2.1 Imprecision on the utilities

Let us start with the first case. Consider a decision problem where we must choose between two alternatives X and Y whose respective utilities depend on the values ω of the states of nature. Assume that we have precise information about the probabilities of these states of nature, so that X and Y can be seen as random variables defined on a probability space $(\Omega, \mathcal{A}, P)$. If we have imprecise knowledge about the utilities $X(\omega)$ associated with the different states of nature, one possible model would be to associate to any $\omega \in \Omega$ a set $\Gamma(\omega)$ that is sure to include the 'true' utility $X(\omega)$. By doing this, we obtain a multi-valued mapping $\Gamma : \Omega \to \mathcal{P}(\Omega)$, and all we know about X is that it is one of the measurable selections of Ω, that were defined in Equation (2.21) by:

$$S(\Gamma) = \{U : \Omega \to \Omega \text{ r.v. } : U(\omega) \in \Gamma(\omega) \text{ for every } \omega \in \Omega\}. \tag{4.11}$$

In this paper, we shall consider only multi-valued mappings satisfying the measurability condition:

$$\Gamma \ (A) := \{\omega \quad \Omega : \Gamma(\omega) \cap A = \quad \} \quad A \quad \text{for any } A \quad A \quad .$$

As we saw in Definition 2.42, these multi-valued mappings are called random sets.

Our comparison of two alternatives with imprecise utilities results thus in the comparison of two random sets Γ_1, Γ_2, that we shall make by means of their respective sets of measurable selections $S(\Gamma_1)$, $S(\Gamma_2)$ determined by Equation (2.21). For simplicity, we shall use the notation $\Gamma_1 \quad \Gamma_2$ instead of $S(\Gamma_1) \quad S(\Gamma_2)$ when no confusion is possible.

Let us begin by studying the comparison of random sets by means of stochastic dominance.

Proposition 4.76 *Let* (Ω, A, P) *be a probability space,* $(\Omega, P(\Omega))$ *a measurable space, with* Ω *a finite subset of* $\mathbb{R}$, *and* Γ_X, Γ_Y *be two random sets. The following equivalences hold:*

(a) $\Gamma_X \quad {}_{FSD_1} \ \Gamma_Y \quad \min \Gamma_X \quad {}_{FSD} \ \max \Gamma_Y$.

(b) $\Gamma_X \quad {}_{FSD_2} \ \Gamma_Y \quad \Gamma_X \quad {}_{FSD_3} \ \Gamma_Y \quad \max \Gamma_X \quad {}_{FSD} \ \max \Gamma_Y$.

(c) $\Gamma_X \quad {}_{FSD_4} \ \Gamma_Y \quad \max \Gamma_X \quad {}_{FSD} \ \min \Gamma_Y$.

(d) $\Gamma_X \quad {}_{FSD_5} \ \Gamma_Y \quad \Gamma_X \quad {}_{FSD_6} \ \Gamma_Y \quad \min \Gamma_X \quad {}_{FSD} \ \min \Gamma_Y$.

Proof The result follows from Proposition 4.19, taking into account that given a random set Γ taking values on a finite space, the lower distribution function associated with its set $S(\Gamma)$ of measurable selections is induced by $\max \Gamma$ and its upper distribution function is induced by $\min \Gamma$. ∎

Moreover, we can characterise the conditions ${}_{FSD_i}$, $i = 1, \ldots, 6$ even for random sets that take values on infinite spaces. To see how this comes out, we shall consider the upper and lower probabilities induced by the random set. Recall that, from Equation (2.22), they are defined by:

$$P \ (A) = P(\{\omega : \Gamma(\omega) \cap A = \quad \}) \text{ and}$$
$$P \ (A) = P(\{\omega : \quad = \Gamma(\omega) \quad A\})$$

for any $A \quad A$. As we have already seen in Equation (2.24), the upper and lower probabilities of a random set constitute upper and lower bounds of the probabilities induced by the measurable selections:

$$P \ (A) \leq P_U(A) \leq P \ (A) \quad U \quad S(\Gamma),$$

and in particular their associated cumulative distributions provide lower and upper bounds of the lower and upper distribution functions associated with $S(\Gamma)$.

We have seen in Theorem 2.46 thatwhen P (A) is attainedbythe probabilities induced by the measurable selections for any element A A, thesupremumand infimum of the integrals of a gamble with resp ect to the measurable selections can be expressed by means of the Cho quet integral of the gamble with resp ect to P and P. This result allows to characterise the imprecise sto chastic dominance b etween random sets by means of the comparison of Cho quet or Aumann integrals. Recall that we have denoted by U the set of increasing and bounded functions $u : [0, 1] \to \mathbb{R}$.

Prop osition 4.77 *Let* (Ω, A, P) *be a probability space. Considerthe measurable space* $([0, 1], \beta_{[0,1]})$ *and let* $\Gamma_X, \Gamma_Y : \Omega \to P([0, 1])$ *betwo randomsets. If for all* A $\beta_{[0,1]}$ *it holds that* $P_X(A) = \max P(\Gamma_X)(A)$ *and* $P_Y(A) = \max P(\Gamma_Y)(A)$*, the fol lowing equiva- lences hold:*

1. Γ_X $_{FSD_1}$ Γ_Y (C) $udP_X \geq (C)$ udP_Y *for every* u U.
2. Γ_X $_{FSD_2}$ Γ_Y (C) $udP_X \geq (C)$ udP_Y *for every* u U.
3. Γ_X $_{FSD_3}$ Γ_Y (C) $udP_X \geq (C)$ udP_X *for every* u U.
4. Γ_X $_{FSD_4}$ Γ_Y (C) $udP_X \geq (C)$ udP_X *for every* u U.
5. Γ_X $_{FSD_5}$ Γ_Y (C) $udP_X \geq (C)$ udP_X *for every* u U.
6. Γ_X $_{FSD_6}$ Γ_Y (C) $udP_X \geq (C)$ udP_Y *for every* u U.

Pro of Consider u U. We deduce from Theorem 2.46 that, under the hyp otheses of the prop osition,

$$(C) \quad udP_X = \sup_{U \quad S(\Gamma_X)} \quad udP_U = \overline{E}_{S(\Gamma_X)}(u) \text{ and}$$

$$(C) \quad udP_X = \inf_{U \quad S(\Gamma_X)} \quad udP_U = \underline{E}_{S(\Gamma_X)}(u)$$

and similarly:

$$(C) \quad udP_Y = \sup_{U \quad S(\Gamma_Y)} \quad udP_U = \overline{E}_{S(\Gamma_Y)}(u) \text{ and}$$

$$(C) \quad udP_Y = \inf_{U \quad S(\Gamma_Y)} \quad udP_U = \underline{E}_{S(\Gamma_Y)}(u)$$

The result followsthen applyingTheorem4.23. ∎

Let us discuss next the comparison of random sets by means of statistical preference. When the utility space Ω is finite, we obtain a result related to Proposition 4.76:

Proposition 4.78 *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space,* $(\Omega, \mathcal{P}(\Omega))$ *a measurable space, with* Ω *finite, and* Γ_X, Γ_Y *be two random sets. The following equivalences hold:*

(a) $\Gamma_X \quad SP_1 \; \Gamma_Y \qquad \min\Gamma_X \quad SP \; \max\Gamma_Y.$

(b) $\Gamma_X \quad SP_2 \; \Gamma_Y \qquad \Gamma_X \quad SP_3 \; \Gamma_Y \qquad \max\Gamma_X \quad SP \; \max\Gamma_Y.$

(c) $\Gamma_X \quad SP_4 \; \Gamma_Y \qquad \max\Gamma_X \quad SP \; \min\Gamma_Y.$

(d) $\Gamma_X \quad SP_5 \; \Gamma_Y \qquad \Gamma_X \quad SP_6 \; \Gamma_Y \qquad \min\Gamma_X \quad SP \; \min\Gamma_Y.$

Proof The result follows from Proposition 4.7, taking into account that statistical preference satisfies the monotonicity condition of Equation (4.2) and that if Γ is a random set taking values on a finite space, then the mappings $\min\Gamma, \max\Gamma$ belong to $S(\Gamma)$. ∎

In particular, we deduce that we can focus on the minimum and maximum measurable selections in order to characterise these extensions of statistical preference.

Corollary 4.79 *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space,* Ω *a finite space and consider two random sets* $\Gamma_X, \Gamma_Y : \Omega \to \mathcal{P}(\Omega)$. *Then for every* $i = 1, \ldots, 6$ *:*

$$\Gamma_X \quad SP_i \; \Gamma_Y \qquad \{\min\Gamma_X, \max\Gamma_X\} \quad SP_i \; \{\min\Gamma_Y, \max\Gamma_Y\}. \tag{4.12}$$

These two results are interesting because random sets taking values on finite spaces are quite common in practice; they have been studied in detail in [59, 127], and one of their most interesting properties is that they constitute equivalent models to belief and plausibility functions [170].

Note that the equivalence in Equation (4.12) does not hold for the relation SP_{mean} defined in Section 4.1.2.

Example 4.80 *Consider the probability space* $(\Omega, \mathcal{A}, P)$ *where* $\Omega = \{\omega_1, \omega_2\}$, $\mathcal{A} = \mathcal{P}(\Omega)$ *and* P *is a probability uniformly distributed on* Ω*, and let* Γ_X *be the random set given by* $\Gamma_X(\omega_1) = \{0, 1\}, \Gamma_X(\omega_2) = \{0, 2, 3, 4\}$*, and let* Γ_Y *be single-valued random set given by* $\Gamma_Y(\omega_1) = \{1\} = \Gamma_Y(\omega_2)$*. Then* $\min\Gamma_X$ *is the constant random variable on* 0*, while* $\max\Gamma_X$ *is given by* $\max\Gamma_X(\omega_1) = 1, \max\Gamma_X(\omega_2) = 4$ *. Hence, if we compare the set* $\{\min\Gamma_X, \max\Gamma_X\}$ *with* Γ_Y *by means of* SP_{mean} *we obtain*

$$\frac{Q(\min\Gamma_X, \Gamma_Y) + Q(\max\Gamma_X, \Gamma_Y)}{2} = \frac{0 + 0.75}{2} = 0.375$$

and thus Γ_Y $_{SP_{mean}}$ $\{\min \Gamma_X, \max \Gamma_X\}$. *Ontheotherhand, thesetofselectionsof* Γ_X *is given by (where aselection* X *is identified with the vect or* $(U(\omega_1), U(\omega_2))$*):*

$$S(\Gamma_X) = \{(0,0), (0,2), (0,3), (0,4), (1,0), (1,2), (1,3), (1,4)\}$$

from which we deduce that Γ_X $_{SP_{mean}}$ Γ_Y.

4.2.2 Imprecision on the beliefs

We next consider the case where we want to cho ose b etween two random variables X and Y defined from Ω to Ω, and there is some uncertainty ab out the probability distribution P ofthe differentstates ofnature ω Ω, that we mo del by means of a set P of probability distributions on Ω. Then we may asso ciate with X aset X of random variables, that corresp ond to the transformations of X underanyoftheprobabilitydistributions in P; and similarly for Y. Weend upthuswith two sets X, Y of random variables, and we should establish metho ds to determine which of these two sets is preferable.

One particular cas e where this situation may arise is in the context of missing data [218]. Wemaydividethevariablesdeterminingthe statesofnatureintwogroups: one for which we have precise information, that we mo del by means of a probability measure P over thedifferent states, and another ab out which are completely ignorant, knowing onlywhichare the different states, but nothing more. Then we may get to the cl assical scenario byfixing the value of the variables in this second group: for each ofthese values the alternatives may be seen as random variables, usingthe probabilitymeasure P to determine the probabilities of the different reward s. Hence, bydoing thiswewould transform the two alternatives X and Y into twosets of alternatives X, Y, considering all the possible values of the variables in the second group.

Inthis situation, we may compare the sets X, Y by means of th e generalisations of statistical preference or sto chastic dominance we have discussed in Section 4.1; however, we argue that other notion s may make more sense in this context. This is because conditions $_1, \ldots, _6$ arebased onconsideringa particularpair (X_1, Y_1) in $X \times Y$ and on comparing X_1 with Y_1 by means of the binary relation . However, any X_1 in X corresp onds to a particular choice of aprobability measure P P, andsimilarly for any Y_1 Y; and if we use an e pistemic interpre tation ofour uncertainty under which only one P P is the 'true' mo del, it makes nosense to compare X_1 and Y_1 based on adifferent distribution. This isparticularlyclearin casewewanttoapply statistical preference, whichisbased oncomparing $P(X > Y)$ with $P(Y > X)$, where P is the initial probability measure.

To make this explicit, in this section we may denote oursets of alternatives by $X := \{(X, P): P \quad P\}$ and $Y := \{(Y, P): P \quad P\}$, meaning that our utilitiesare precise (and are determinedby the variables X and Y, resp ectively), while our beliefs are imprecise and are mo delled by the set P. Toavoid confusions, we will now write

$X \ ^{P}\ Y$ toexpress that X is preferred to Y whenweconsidertheprobability measure P in the initial probability space. Then we can establish the following definitions:

Definition 4.81 *Let be a binary relationon random variables. Wesay that:*

- X *is* strongly P preferred *to* Y*, anddenote it* $X \ {}^{P}_{s}\ Y$*, when* $X \ ^{P}\ Y$ *for every* $P \ P$ *;*
- X *is* weakly P preferred*, anddenote it* $X \ {}^{P}_{w}\ Y$*, to* Y *when* $X \ ^{P}\ Y$ *for some* $P \ P$ *.*

Obviously, th e strong preference implies the weak one.Tosee thattheyare notequiva-lent, consider the follow ing simple example:

Example 4.82 *Let be the binary relation associated with statistical preference and consider the variables* X, Y *that represent theresults ofthe dices* A *and* B*, respectively, in Example 3.83. If weconsider the uniform distribution* P_1 *in all the dieoutcomes, we obtain* $Q(X, Y) = \frac{5}{9}$*, so that* $X \ {}^{P_1}_{SP}\ Y$*; if we take instead the uniform distribution* P_2 *on* $\{1, 2, 3\}$*, then* $Q(X, Y) = \frac{1}{9}$*, and asa consequence* $Y \ {}^{P_2}_{SP}\ X$*. Hence,* X *is weakly* $\{P_1, P_2\}$ *statistical ly preferred to* Y*, but not strongly so.*

With resp ect to the notions established in Section 4.1, it is not difficult to establish the following res ult. Its pro of is immediate, and therefore omitted.

Prop osition 4.83 *Let* X, Y *be the setsof alternatives consideredabove, and let bea binary relation. Then*

$$X \ _1 Y \quad X \ {}^{P}_{s}\ Y \quad X \ {}^{P}_{w}\ Y \quad X \ _4 Y.$$

To see that the converse implicationsdo not hold, consider the followingexample:

Example 4.84 *Consider* $\Omega = \{\omega_1, \omega_2, \omega_3\}$*, the set of probabilities*

$$P := \{P : P(\omega_1) > P(\omega_2), P(\omega_2) \quad [0, 0.2]\}$$

and the alternatives X, Y *given by*

	ω_1	ω_2	ω_3
X	1	0	1
Y	0	1	1

If weconsider thesets $X = \{(X, P) : P \ P\}$ *and* $Y = \{(Y, P) : P \ P\}$ *andwe compare them by means of stochastic dominance, itis clearthat* $X \ {}^{P}_{s}\ Y$*; however, itdoes not* *hold that* X $_{FSD_1}$ Y*: if we consider* $P_1 := (0.3, 0.2, 0.5)$ *and* $P_2 := (0.1, 0, 0.9)$*, it holds that* (Y, P_2) $_{FSD}$ (X, P_1).

Moreover, in this example we also have that X *is strictly weakly* P*-preferred to* Y *while* $X \equiv_{FSD_4} Y$.

Remark 4.85 *If the binary relation we start with is complete, so is the weak* P*-preference. In that case, we obtain that* X ^P_w Y *implies that* X ^P_s Y*, because if* X ^P_w Y *we must have that* (X, P) (Y, P) *for every* P P.

Moreover, when $X \equiv^P_w Y$*, we may have strict preference, indifference or incomparability with respect to strong* P*-preference.*

In what follows, we study in some detail the notions of weak and strong preference for particular choices of the binary relation . If corresponds to expected utility, strong preference of X over Y means that X is preferred to Y with respect to all the probability measures P in P, and then it is related to the idea of *maximality* [205]; on the other hand, weak preference means that X is preferred to Y (i.e., it is the optimal alternative) with respect to some of the elements of P; this idea is close to the criterion of E-admissibility [107]. See also Remark 4.13 and [43, Section 3.2].

When is the binary relation associated with stochastic dominance, we obtain the following.

Proposition 4.86 *Consider a set* P *of probability measures on* Ω*, and let* X, Y *be two real-valued random variables on* Ω*. Let us define the sets* $F_X := \{F_X^P : P \ P\}$ *and* $F_Y := \{F_Y^P : P \ P\}$.

1. $\overline{F}_X \leq \underline{F}_Y$ X *is strongly* P*-preferred to* Y *with respect to stochastic dominance.*
2. X *is weakly* P*-preferred to* Y *with respect to stochastic dominance* $\underline{F}_X \leq \overline{F}_Y$.

Proof Assume that $\overline{F}_X \leq \underline{F}_Y$. Then, for any P P it holds that:

$$F_X^P \leq \overline{F}_X \leq \underline{F}_Y \leq F_Y^P.$$

Then, X is strongly P-preferred to Y with respect to first degree stochastic dominance.

Now, assume that X is weakly P-preferred to Y with respect to first degree stochastic dominance. Then there exists P P such that $F_X^P \leq F_Y^P$. Then, in particular, X $_{FSD_4}$ Y, and by Proposition 4.19 we deduce that $\underline{F}_X \leq \overline{F}_Y$. ■

Note that this result could also be derived from Propositions 4.19 and 4.83.

Finally, when corresponds to statistical preference, we can apply Remark 4.85, because is a complete relation. In addition, we can establish the following result:

Proposition 4.87 *Consider a set $\mathcal{P}$ of probability measures, and let $\underline{P}, \overline{P}$ denote its lower and upper envelopes, given by Equation* (2.18). *Let X, Y be two real-valued random variables on Ω, and let $u = I_{(0,+\infty)} - I_{(-\infty,0)}$.*

1. *X is strongly $\mathcal{P}$ statistically preferred to Y $\Leftrightarrow$ $\underline{P}(u(X-Y)) \geq 0$.*
2. *X is weakly $\mathcal{P}$ statistically preferred to Y $\Leftarrow$ $\overline{P}(u(X-Y)) \geq 0$. The converse holds if $\mathcal{P} = \mathcal{M}(\underline{P})$.*

Proof The result follows simply by considering that if X, Y are random variables on a probability space $(\Omega, \mathcal{A}, P)$, then, by applying Equation (3.1), $X \succeq_{SP}^{P} Y$ if and only if $P(u(X-Y)) \geq 0$, where we also use P to denote the expectation operator associated with the probability measure P.

To see that the converse of the second statement holds when $\mathcal{P} = \mathcal{M}(\underline{P})$, note that the upper envelope $\overline{P}$ of $\mathcal{P}$ is a coherent lower prevision. From [205, Section 3.3.3], given the bounded random variable $u(X-Y)$ there exists a probability P in $\mathcal{M}(\underline{P})$ such that $P(u(X-Y)) = \overline{P}(u(X-Y))$. ■

The above result can be related to the lower median, as in [46, 148]. For this, let us define the *lower median* of $X - Y$ by the credal set $\mathcal{M}(\underline{P})$ by

$$\underline{Me}(X-Y) := \inf\{Me_P(X-Y) : P \in \mathcal{M}(\underline{P})\},$$

and its *upper median* by

$$\overline{Me}(X-Y) := \sup\{Me_P(X-Y) : P \in \mathcal{M}(\underline{P})\},$$

where $Me_P(X-Y)$ denotes the median of $X-Y$ when P is the probability of the initial space.

Then, we deduce from Proposition 4.64 that

$$\underline{Me}(X-Y) > 0 \Rightarrow X \succeq_{SP,s}^{\mathcal{M}(\underline{P})} Y \Rightarrow \underline{Me}(X-Y) \geq 0,$$

and that

$$\overline{Me}(X-Y) > 0 \Rightarrow X \succeq_{SP,w}^{\mathcal{M}(\underline{P})} Y \Rightarrow \overline{Me}(X-Y) \geq 0.$$

A related result was established in [46, Proposition 4], by means of a slightly different definition of median. See also Proposition 4.64, and [83, 164] for approaches based on the expected utility model.

4.3 Modelling the joint distribution in an imprecise framework

Statistical preference is an stochastic order that depends on the joint distribution of the random variables. This joint distribution function can be determined, according to Sklar's Theorem (Theorem 2.27), from the marginals by means of a copula. In the imprecise context we are dealing with in this chapter, there may be imprecision either in the marginal distribution functions or in the copula that links the marginals. In the former case, we can model the lack of information by means of p-boxes, and in the second one the should consider a set of possible copulas. In both situations we shall obtain a set of bivariate distribution functions.

In order to determine the mathematical model for this situation, we shall consider two steps: on the one hand, we shall study how to model sets of bivariate distribution functions, since the lower and upper bounds are not, in general, distribution functions. To deal with this problem, we shall extend the notion of p-box when considering bivariate distribution functions, and we will investigate under which conditions such bivariate p-box can define a coherent lower probability. Then, we shall consider two marginal imprecise distribution functions and we will try to build from them a joint distribution. In this context, the main result is to extend Sklar's Theorem to an imprecise framework; we shall also study the application of these results can be applied into bivariate stochastic orders.

4.3.1 Bivariate distribution with imprecision

Bivariate p-boxes

Let Ω_1, Ω_2 be two totally ordered spaces. As in [198], we assume without loss of generality that both have a maximum element, that we denote respectively by x^*, y^*. Note that this is trivial in the case of finite spaces.

We start by introducing standardized functions and bivariate distribution functions.

Definition 4.88 *Consider two ordered spaces Ω_1, Ω_2. A map $F: \Omega_1 \times \Omega_2 \to [0,1]$ is called* standardized *when it is component-wise increasing and $F(x^*, y^*) = 1$. It is called a* distribution function *when moreover it satisfies the* rectangle inequality*:*

$$(\mathbf{RI}): \quad F(x_2,y_2) + F(x_1,y_1) - F(x_1,y_2) - F(x_2,y_1) \geq 0$$

for every $x_1, x_2 \in \Omega_1$ and $y_1, y_2 \in \Omega_2$ such that $x_1 \leq x_2$ and $y_1 \leq y_2$.

Here, and in what follows, we shall make an assumption of *logical independence*, meaning that we consider all values in the product space $\Omega_1 \times \Omega_2$ to be possible.

The rectangle inequality is equivalent to monotonicity in the univariate case, so in that case a distribution function is simply an increasing and normalized function $F: X \to [0,1]$. Moreover, a lower envelope of univariate distribution functions is again a distribution function, by Proposition 2.34. Unfortunately, the situation is not as clear cut in the bivariate case: the envelopes of a set of distribution functions are standardized maps, but not necessarily distribution functions.

Proposition 4.89 *Let Ω_1 and Ω_2 be two ordered spaces and $\mathcal{F}$ be a family of distribution functions $F : \Omega_1 \times \Omega_2 \to [0,1]$. Their lower and upper envelopes $\underline{F}, \overline{F} : \Omega_1 \times \Omega_2 \to [0,1]$ given by*

$$\underline{F}(x, y) = \inf_{F \in \mathcal{F}} F(x, y) \text{ and } \overline{F}(x, y) = \sup_{F \in \mathcal{F}} F(x, y)$$

for every $x \in \Omega_1, y \in \Omega_2$, are standardized maps.

Proof It suffices to take into account that the monotonicity and normalization properties are preserved by lower and upper envelopes. ∎

To see that these envelopes are not necessarily distribution functions, consider the following example:

Example 4.90 *Take $\Omega_1 = \Omega_2 = \{a, b, c\}$, with $a < b < c$ and let F_1, F_2 be the distribution functions determined by the following joint probability measures:*

X_1, Y_1	a	b	c
a	0.1	0.1	0
b	0.4	0.1	0
c	0	0	0.3

X_2, Y_2	a	b	c
a	0.4	0	0.2
b	0.1	0	0
c	0.1	0	0.2

Then F_1 and F_2 are given by:

	(a, a)	(a, b)	(a, c)	(b, a)	(b, b)	(b, c)	(c, a)	(c, b)	(c, c)
F_1	0.1	0.2	0.2	0.5	0.7	0.7	0.5	0.7	1
F_2	0.4	0.4	0.6	0.5	0.5	0.7	0.6	0.6	1

and their lower and upper envelopes are given by:

	(a, a)	(a, b)	(a, c)	(b, a)	(b, b)	(b, c)	(c, a)	(c, b)	(c, c)
$\underline{F}$	0.1	0.2	0.2	0.5	0.5	0.7	0.5	0.6	1
$\overline{F}$	0.4	0.4	0.6	0.5	0.7	0.7	0.6	0.7	1

Then

$$\underline{F}(b,b) + \underline{F}(a,a) - \underline{F}(a, b) - \underline{F}(b, a) = 0.5 + 0.1 - 0.2 - 0.5 = -0.1 < 0$$

and

$$\overline{F}(b,c) + \overline{F}(a,b) - \overline{F}(a, c) - \overline{F}(b, b) = 0.7 + 0.4 - 0.6 - 0.7 = -0.2 < 0.$$

As a consequence, neither $\underline{F}$ nor $\overline{F}$ are distribution functions.

Taking this result into account, we give the following definition:

Definition 4.91 *Consider two ordered spaces* Ω_1, Ω_2*, and let* $\underline{F}, \overline{F} : \Omega_1 \times \Omega_2 \to [0,1]$ *be two standardized functions satisfying* $\underline{F}(x,y) \leq \overline{F}(x,y)$ *for every* $x \ \ \Omega_1, y \ \ \Omega_2$*. Then the pair* $(\underline{F}, \overline{F})$ *is called a* bivariate p-box*.*

Proposition 4.89 shows that bivariate p-boxes can be obtained in particular by means of a set of distribution functions, taking lower and upper envelopes. However, not all bivariate p-boxes are of this type: if we consider for instance a map $\underline{F} = \overline{F}$ that is standardized but not a distribution function, then there is no bivariate distribution function between $\underline{F}$ and $\overline{F}$, and as a consequence these cannot be obtained as envelopes of a set of distribution functions. Our next paragraph will deepen into this matter, by means of the notion of coherence of lower probabilities. In particular, we shall investigate how Theorem 2.35 could be extended to bivariate p-boxes.

Lower probabilities and p-boxes

In order to define a lower probability from a bivariate p-box, let us now introduce a notation similar to the one of Section 2.2.1.

Consider two ordered spaces Ω_1, Ω_2, and let $(\underline{F}, \overline{F})$ be a bivariate p-box on $\Omega_1 \times \Omega_2$. Denote

$$A_{(x,y)} := \{(x', y') \ \ \Omega_1 \times \Omega_2 : x' \leq x, y' \leq y\},$$

and let us define

$$K_1 := \{A_{(x,y)} : x \ \ \Omega_1, y \ \ \Omega_2\} \text{ and } K_2 := \{A^c_{(x,y)} : x \ \ \Omega_1, y \ \ \Omega_2\}.$$

The maps $\underline{F}$ and $\overline{F}$ can be used to define the lower probabilities $\underline{P}_{\underline{F}} : K_1 \to \mathbb{R}$ and $\underline{P}_{\overline{F}} : K_2 \to \mathbb{R}$ by:

$$\underline{P}_{\underline{F}}(A_{(x,y)}) = \underline{F}(x,y) \quad \text{and} \quad \underline{P}_{\overline{F}}(A^c_{(x,y)}) = 1 - \overline{F}(x,y). \tag{4.13}$$

Define now $K := K_1 \ \ K_2$; note that $A_{(x^*, y^*)} = \Omega_1 \times \Omega_2$, where x^*, y^* are the maximum of Ω_1 and Ω_2, respectively. Thus, both $\Omega_1 \times \Omega_2$ and belong to K.

Definition 4.92 *The* lower probability induced by $(\underline{F}, \overline{F})$ *is the map* $\underline{P}_{(\underline{F},\overline{F})} : K \to [0,1]$ *given by:*

$$\underline{P}_{(\underline{F},\overline{F})}(A_{(x,y)}) = \underline{F}(x,y), \quad \underline{P}_{(\underline{F},\overline{F})}(A^c_{(x,y)}) = 1 - \overline{F}(x,y) \tag{4.14}$$

for every $x \ \ \Omega_1, y \ \ \Omega_2$*.*

Note that $\underline{P}_{(\underline{F},\overline{F})}(\Omega_1 \times \Omega_2) = 1$ and $\underline{P}_{(\underline{F},\overline{F})}(\ ) = 0$ because $\underline{F}$ and $\overline{F}$ are standardized.

In this section, we are going to study which properties of the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ can be characterised in terms of the lower and upper distribution functions $\underline{F}$ and $\overline{F}$.

Avoiding sure loss We begin with the property of avoiding sure loss. Recall that, as we saw in Definition 2.29, a lower probability $\underline{P}$ with domain $\mathcal{K} \subseteq \mathcal{P}(\Omega_1 \times \Omega_2)$ avoids sure loss if and only if there is a finitely additive probability $P: \mathcal{P}(\Omega_1 \times \Omega_2) \to [0,1]$ that dominates $\underline{P}$ on its domain. This is a consequence of [205, Corollary 3.2.3 and Theorem 3.3.3].

Proposition 4.93 *The lower probability $\underline{P}_{(\underline{F},\overline{F})}$ induced by the bivariate p-box $(\underline{F},\overline{F})$ by means of Equation (4.14) avoids sure loss if and only if there is a distribution function $F: \Omega_1 \times \Omega_2 \to [0,1]$ satisfying $\underline{F} \le F \le \overline{F}$.*

Proof We begin with the direct implication. Assume that $\underline{P}_{(\underline{F},\overline{F})}$ avoids sure loss. Then, there exists a finitely additive probability $P: \mathcal{P}(\Omega_1 \times \Omega_2) \to [0,1]$ such that $P(A) \ge \underline{P}_{(\underline{F},\overline{F})}(A)$ for every $A \in \mathcal{K}$. Let us define the map $F_P: \Omega_1 \times \Omega_2 \to [0,1]$ by $F_P(x,y) = P(A_{(x,y)})$. Then F_P is a distribution function that is bounded between $\underline{F}$ and $\overline{F}$:

- Consider $x_1, x_2 \in \Omega_1$ and $y_1, y_2 \in \Omega_2$ such that $x_1 \le x_2, y_1 \le y_2$. Then:
$$F_P(x_1,y_1) = P(A_{(x_1,y_1)}) \le P(A_{(x_2,y_2)}) = F_P(x_2,y_2)$$
because P is monotone.
- $F_P(x^*,y^*) = P(A_{(x^*,y^*)}) = P(\Omega_1 \times \Omega_2) = 1$.
- Consider $x_1, x_2 \in \Omega_1$ and $y_1, y_2 \in \Omega_2$ such that $x_1 \le x_2, y_1 \le y_2$. Then it holds that
$$\begin{aligned} &F_P(x_1,y_1) + F_P(x_2,y_2) - F_P(x_1,y_2) - F_P(x_2,y_1) \\ &= P(A_{(x_1,y_1)}) + P(A_{(x_2,y_2)}) - P(A_{(x_1,y_2)}) - P(A_{(x_2,y_1)}) \\ &= P(\{(x,y) \in \Omega_1 \times \Omega_2 : x_1 < x \le x_2, y_1 < y \le y_2\}) \ge 0. \end{aligned}$$
- For every $x \in \Omega_1, y \in \Omega_2$,
$$F_P(x,y) = P(A_{(x,y)}) \ge \underline{P}_{(\underline{F},\overline{F})}(A_{(x,y)}) = \underline{F}(x,y),$$
and on the other hand,
$$\begin{aligned} F_P(x,y) &= P(A_{(x,y)}) = 1 - P(A^c_{(x,y)}) \\ &\le 1 - \underline{P}_{(\underline{F},\overline{F})}(A^c_{(x,y)}) = 1 - (1 - \overline{F}(x,y)) = \overline{F}(x,y). \end{aligned}$$

Conversely, assume that $F: \Omega_1 \times \Omega_2 \to [0,1]$ is a distribution function that lies between $\underline{F}$ and $\overline{F}$, and let us define the finitely additive probability P_F on the field generated by $\mathcal{K}$ by means of
$$\begin{aligned} &P_F(\{(x,y) \in \Omega_1 \times \Omega_2 : x_1 < x \le x_2, y_1 < y \le y_2\}) \\ &= F_P(x_1,y_1) + F_P(x_2,y_2) - F_P(x_1,y_2) - F_P(x_2,y_1) \ge 0. \end{aligned} \tag{4.15}$$

Then it follows that $P_F(A_{(x,y)}) = F(x,y) \geq \underline{F}(x,y) = \underline{P}_{(\underline{F},\overline{F})}(A_{(x,y)})$ and moreover $P_F(A^c_{(x,y)}) = 1 - F(x,y) \geq 1 - \underline{F}(x,y) = \underline{P}_{(\underline{F},\overline{F})}(A^c_{(x,y)})$.

Since any finitely additive probability on a field of events has a finitely additive extension to $\mathcal{P}(\Omega_1 \times \Omega_2)$, we deduce that there is a finitely additive probability that dominates $\underline{P}_{(\underline{F},\overline{F})}$, and as a consequence this lower probability avoids sure loss. ∎

This result allows us to focus on the lower and upper distributions of the p-box, that shall simplify search for for necessary and sufficient conditions. We shall say that $(\underline{F},\overline{F})$ *avoids sure loss* when the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ it induces by means of Equation (4.14) does. Our next result gives a necessary condition:

Proposition 4.94 *If $(\underline{F},\overline{F})$ avoids sure loss, then for every $x_1, x_2 \in \Omega_1$ and $y_1, y_2 \in \Omega_2$ such that $x_1 \leq x_2$ and $y_1 \leq y_2$ it holds that*

$$(\textbf{I}-\textbf{RI0}): \quad \overline{F}(x_2,y_2) + \overline{F}(x_1,y_1) - \underline{F}(x_1,y_2) - \underline{F}(x_2,y_1) \geq 0.$$

Proof Assume that $(\underline{F},\overline{F})$ avoids sure loss. By Proposition 4.93, there is a distribution function F bounded by $\underline{F},\overline{F}$. Given $x_1, x_2 \in \Omega_1$ and $y_1, y_2 \in \Omega_2$ such that $x_1 \leq x_2$ and $y_1 \leq y_2$, it follows from (RI) that

$$\begin{aligned} 0 &\leq F(x_2,y_2) + F(x_1,y_1) - F(x_1,y_2) - F(x_2,y_1) \\ &\leq \overline{F}(x_2,y_2) + \overline{F}(x_1,y_1) - \underline{F}(x_1,y_2) - \underline{F}(x_2,y_1), \end{aligned}$$

where the second inequality follows from $\underline{F} \leq F \leq \overline{F}$. ∎

Let us show that this necessary condition is not sufficient in general:

Example 4.95 *Consider $\Omega_1 = \Omega_2 = \{a,b,c\}$, with $a < b < c$ and let $\underline{F}$ and $\overline{F}$ be given by:*

	(a,a)	(a,b)	(a,c)	(b,a)	(b,b)	(b,c)	(c,a)	(c,b)	(c,c)
$\underline{F}$	0	0.65	0.7	0.2	0.8	0.8	0.35	0.9	1
$\overline{F}$	0.1	0.7	0.7	0.25	0.8	0.8	0.4	0.9	1

It is immediate to check that both maps are standardized and that together they satisfy (I-RI0). However, $(\underline{F},\overline{F})$ does not avoid sure loss: from Proposition 4.93, it suffices to show that there is no distribution function F bounded by $\underline{F}(x,y)$ and $\overline{F}(x,y)$ for every $x,y \in \{a,b,c\}$. To see that this is indeed the case, note that any distribution function $F \in (\underline{F},\overline{F})$ should satisfy

$$F(a,c) = 0.7,\ F(b,b) = 0.8,\ F(b,c) = 0.8,\ F(c,b) = 0.9 \text{ and } F(c,c) = 1.$$

By (RI) to $(x_1,y_1) = (a,b)$ and $(x_2,y_2) = (b,c)$, we deduce that $F(a,b) = 0.7$, and then applying again the rectangle inequality we deduce that

$$F(b,b) + F(a,a) - F(a,b) - F(b,a) = 0.8 + F(a,a) - 0.7 - F(b,a) \geq 0$$

if and only if $F(a,a) + 0.1 \geq F(b,a)$ *whence* $F(a,a) = 0.1$ *and* $F(b,a) = 0.2$. *If we now apply (RI) to* $(x_1,y_1) = (b,a)$ *and* $(x_2,y_2) = (c,b)$, *we deduce that*

$$F(c,b) + F(b,a) - F(b,b) - F(c,a) = 0.9 + 0.2 - 0.8 - F(c,a) \geq 0$$

if and only if $F(c,a) \leq 0.3$. *But on the other hand we must have* $F(c,a) \geq \underline{F}(c,a) = 0.35$, *a contradiction. Hence,* $(\underline{F},\overline{F})$ *does not avoid sure loss.*

However, (I-RI0) is a necessary and sufficient condition when both Ω_1,Ω_2 are binary spaces.

Proposition 4.96 *Assume that both* $\Omega_1 = \{x_1,x_2\}$ *and* $\Omega_2 = \{y_1,y_2\}$ *are binary spaces such that* $x_1 \leq x_2$ *and* $y_1 \leq y_2$, *and let* $(\underline{F},\overline{F})$ *be a bivariate p-box on* $\Omega_1 \times \Omega_2$. *Then the following are equivalent:*

1. $(\underline{F},\overline{F})$ *avoids sure loss.*
2. $\overline{F}(x_2,y_2) + \overline{F}(x_1,y_1) - \underline{F}(x_1,y_2) - \underline{F}(x_2,y_1) \geq 0$ *for all* $x_1,x_2 \in \Omega_1, y_1,y_2 \in \Omega_2$.
3. $\underline{F}(x_2,y_2) + \overline{F}(x_1,y_1) - \underline{F}(x_1,y_2) - \underline{F}(x_2,y_1) \geq 0$ *for all* $x_1,x_2 \in \Omega_1, y_1,y_2 \in \Omega_2$.

Proof The first statement implies the second from Proposition 4.94. To see that the second implies the third note that, since $\underline{F}$ and $\overline{F}$ are standardized maps, it holds that $\underline{F}(x_2,y_2) = \overline{F}(x_2,y_2) = 1$.

To see that the third statement implies the first, let us consider $F : \Omega_1 \times \Omega_2 \to [0,1]$ given by

$$\begin{aligned}
F(x_1,y_1) &= \overline{F}(x_1,y_1)\\
F(x_1,y_2) &= \max\{\overline{F}(x_1,y_1), \underline{F}(x_1,y_2)\}\\
F(x_2,y_1) &= \max\{\overline{F}(x_1,y_1), \underline{F}(x_2,y_1)\}\\
F(x_2,y_2) &= 1.
\end{aligned}$$

By construction, F is a standardized map and it is bounded by $\underline{F},\overline{F}$. To see that it indeed is a distribution function, note that if either $F(x_1,y_2)$ or $F(x_2,y_1)$ is equal to $F(x_1,y_1) = F(x_1,y_1)$, then it follows from the monotonicity of $\underline{F},\overline{F}$ that

$$F(x_2,y_2) + F(x_1,y_1) - F(x_1,y_2) - F(x_2,y_1) \geq 0;$$

and if $F(x_1,y_2) = \underline{F}(x_1,y_2)$ and $F(x_2,y_1) = \underline{F}(x_2,y_1)$, then

$$\begin{aligned}
&F(x_2,y_2) + F(x_1,y_1) - F(x_1,y_2) - F(x_2,y_1)\\
&\quad = \underline{F}(x_2,y_2) + \overline{F}(x_1,y_1) - \underline{F}(x_1,y_2) - \underline{F}(x_2,y_1) \geq 0.
\end{aligned}$$

Coherence Let us turn now to coherence, where we shall see that Theorem 2.35 does not extend immediately to the bivariate case. We begin by establishing a result related to Proposition 4.93:

Proposition 4.97 *The lower probability $\underline{P}_{(\underline{F},\overline{F})}$ induced by the bivariate p-box $(\underline{F},\overline{F})$ is coherent if and only if $\underline{F}$ and $\overline{F}$ are the lower and the upper envelopes of the set*

$$\{F : \Omega_1 \times \Omega_2 \to [0,1] \text{ distribution function} : \underline{F} \leq F \leq \overline{F}\},$$

respectively.

Proof We begin with the direct implication. If $\underline{P}_{(\underline{F},\overline{F})}$ is coherent, then for any $x \in \Omega_1$ and $y \in \Omega_2$ there is some probability $P \geq \underline{P}_{(\underline{F},\overline{F})}$ such that $P(A_{(x,y)}) = \underline{P}_{(\underline{F},\overline{F})}(A_{(x,y)})$. Consider the function $F_P : \Omega_1 \times \Omega_2 \to [0,1]$ defined by $F_P(x,y) = P(A_{(x,y)})$ for every $(x,y) \in \Omega_1 \times \Omega_2$. Reasoning as in the proof of Proposition 4.93, we deduce that F_P is a distribution function that belongs to $(\underline{F},\overline{F})$. Moreover, by construction:

$$F_P(x,y) = P(A_{(x,y)}) = \underline{P}_{(\underline{F},\overline{F})}(A_{(x,y)}) = \underline{F}(x,y).$$

Similarly, there exists some $P \geq \underline{P}_{(\underline{F},\overline{F})}$ such that

$$P(A^c_{(x,y)}) = \underline{P}_{(\underline{F},\overline{F})}(A^c_{(x,y)}).$$

Let $F_P : \Omega_1 \times \Omega_2 \to [0,1]$ be given by $F_P(x,y) = P(A_{(x,y)})$ for every $(x,y) \in \Omega_1 \times \Omega_2$. Reasoning as in the proof of Proposition 4.93, we deduce that F_P is a distribution function that belongs to $(\underline{F},\overline{F})$. Moreover, by construction:

$$1 - F_P(x,y) = 1 - P(A_{(x,y)}) = P(A^c_{(x,y)}) = \underline{P}_{(\underline{F},\overline{F})}(A^c_{(x,y)}) = 1 - \overline{F}(x,y),$$

whence $F_P(x,y) = \overline{F}(x,y)$.

Conversely, fix $(x,y) \in \Omega_1 \times \Omega_2$ and let F_1, F_2 be distribution functions in $(\underline{F},\overline{F})$ such that $F_1(x,y) = \underline{F}(x,y)$ and $F_2(x,y) = \overline{F}(x,y)$. Let P_1, P_2 be the finitely additive probabilities they induce in $\mathcal{K}$ by means of Equation (4.15). Then it follows from the proof of Proposition 4.93 that P_1, P_2 dominate $\underline{P}_{(\underline{F},\overline{F})}$, and moreover

$$P_1(A_{(x,y)}) = F_1(x,y) = \underline{F}(x,y) = \underline{P}_{(\underline{F},\overline{F})}(A_{x,y}) \text{ and}$$
$$P_2(A^c_{(x,y)}) = 1 - P_2(A_{(x,y)}) = 1 - F_2(x,y) = 1 - \overline{F}(x,y) = \underline{P}_{(\underline{F},\overline{F})}(A^c_{x,y})$$

Since P_1, P_2 have finitely additive extensions to $\mathcal{P}(\Omega_1 \times \Omega_2)$, we deduce from this that $\underline{P}_{(\underline{F},\overline{F})}$ is coherent. ∎

We shall call the bivariate p-box $(\underline{F},\overline{F})$ *coherent* when its associated lower probability is. One interesting difference with the univariate case is that $\underline{F},\overline{F}$ need not be distribution functions for $(\underline{F}, \overline{F})$ to be coherent (although if $\underline{F}, \overline{F}$ are distribution functions then trivially $(\underline{F}, \overline{F})$ is coherent by Proposition 4.97). This can be seen for instance with Example 4.90, where the lower envelope of a set of distribution functions (which determines the lower distribution function of a coherent p-box) is not a distribution function itself.

Out next result uses properties (2.11)–(2.15) of coherent lower probabilities to obtain four imprecise-versions of the rectangle inequality that, as we shall see, will play an important role.

Proposition 4.98 *Let $(\underline{F}, \overline{F})$ be a bivariate p-box on $\Omega_1 \times \Omega_2$. If it is coherent, then the following conditions hold for every $x_1, x_2 \in \Omega_1$ and $y_1, y_2 \in \Omega_2$ such that $x_1 \leq x_2$ and $y_1 \leq y_2$:*

$$(\mathbf{I-RI1}): \quad \underline{F}(x_2,y_2) + \overline{F}(x_1,y_1) - \underline{F}(x_1,y_2) - \underline{F}(x_2,y_1) \geq 0.$$
$$(\mathbf{I-RI2}): \quad \overline{F}(x_2,y_2) + \underline{F}(x_1,y_1) - \underline{F}(x_1,y_2) - \underline{F}(x_2,y_1) \geq 0.$$
$$(\mathbf{I-RI3}): \quad \overline{F}(x_2,y_2) + \overline{F}(x_1,y_1) - \overline{F}(x_1,y_2) - \underline{F}(x_2,y_1) \geq 0.$$
$$(\mathbf{I-RI4}): \quad \overline{F}(x_2,y_2) + \overline{F}(x_1,y_1) - \underline{F}(x_1,y_2) - \overline{F}(x_2,y_1) \geq 0.$$

Proof Consider (x_1,y_1) and (x_2,y_2) in $\Omega_1 \times \Omega_2$ such that $x_1 \leq x_2$ and $y_1 \leq y_2$. Let $\underline{P}_{(\underline{F},\overline{F})}$ be the lower probability induced by $(\underline{F}, \overline{F})$ by means of Equation (4.14). It is coherent by Proposition 4.97.

Then, by Equations (2.11) and (2.13), it holds that:

$$\begin{aligned}
\underline{P}(A_{(x_2,y_2)}) &\geq \underline{P}(A_{(x_1,y_2)} \cup A_{(x_2,y_1)}) + \underline{P}(A_{(x_2,y_2)} \setminus (A_{(x_1,y_2)} \cup A_{(x_2,y_1)})) \\
&\geq \underline{P}(A_{(x_1,y_2)}) + \underline{P}(A_{(x_2,y_1)}) - \overline{P}(A_{(x_1,y_2)} \cap A_{(x_2,y_1)}) \\
&\quad + \underline{P}(A_{(x_2,y_2)} \setminus (A_{(x_1,y_2)} \cup A_{(x_2,y_1)})).
\end{aligned}$$

Thus:

$$\begin{aligned}
\underline{P}(A_{(x_2,y_2)}) - \underline{P}(A_{(x_1,y_2)}) - \underline{P}(A_{(x_2,y_1)}) + \overline{P}(A_{(x_1,y_2)} \cap A_{(x_2,y_1)}) \\
\geq \underline{P}(A_{(x_2,y_2)} \setminus (A_{(x_1,y_2)} \cup A_{(x_2,y_1)})) \geq 0.
\end{aligned}$$

If we write the previous equation in terms of the maps $\underline{F}, \overline{F}$, we obtain that:

$$\underline{F}(x_2,y_2) - \underline{F}(x_1,y_2) - \underline{F}(x_2,y_1) + \overline{F}(x_1,y_1) \geq 0.$$

On the other hand, applying Equations (2.12) and (2.14)

$$\begin{aligned}
\overline{P}(A_{(x_2,y_2)}) &\geq \overline{P}(A_{(x_1,y_2)} \cup A_{(x_2,y_1)}) + \underline{P}(A_{(x_2,y_2)} \setminus (A_{(x_1,y_2)} \cup A_{(x_2,y_1)})) \\
&\geq \underline{P}(A_{(x_1,y_2)}) + \underline{P}(A_{(x_2,y_1)}) - \underline{P}(A_{(x_1,y_2)} \cap A_{(x_2,y_1)}) \\
&\quad + \underline{P}(A_{(x_2,y_2)} \setminus (A_{(x_1,y_2)} \cup A_{(x_2,y_1)})).
\end{aligned}$$

Then:

$$\overline{P}(A_{(x_2,y_2)}) + \underline{P}(A_{(x_1,y_2)} \cap A_{(x_2,y_1)}) - \underline{P}(A_{(x_1,y_2)}) - \underline{P}(A_{(x_2,y_1)}) \geq \underline{P}(A_{(x_2,y_2)} \setminus (A_{(x_1,y_2)} \cup A_{(x_2,y_1)})) \geq 0.$$

In terms of $\underline{F},\overline{F}$, this means that

$$\overline{F}(x_2,y_2) + \underline{F}(x_1,y_1) - \underline{F}(x_1,y_2) - \underline{F}(x_2,y_1) \geq 0.$$

Analogously, by Equation (2.12)

$$\overline{P}(A_{(x_2,y_2)}) \geq \overline{P}(A_{(x_1,y_2)} \cup A_{(x_2,y_1)}) + \underline{P}(A_{(x_2,y_2)} \setminus (A_{(x_1,y_2)} \cup A_{(x_2,y_1)}))$$

and, from Equation (2.15), this is greater than or equal to both

$$\underline{P}(A_{(x_2,y_2)} \setminus (A_{(x_1,y_2)} \cup A_{(x_2,y_1)})) + \underline{P}(A_{(x_1,y_2)}) + \overline{P}(A_{(x_2,y_1)}) - \overline{P}(A_{(x_1,y_2)} \cap A_{(x_2,y_1)})$$

and

$$\underline{P}(A_{(x_2,y_2)} \setminus (A_{(x_1,y_2)} \cup A_{(x_2,y_1)})) + \overline{P}(A_{(x_1,y_2)}) + \underline{P}(A_{(x_2,y_1)}) - \overline{P}(A_{(x_1,y_2)} \cap A_{(x_2,y_1)}).$$

Then:

$$\begin{aligned} 0 \leq\ & \underline{P}(A_{(x_2,y_2)} \setminus (A_{(x_1,y_2)} \cup A_{(x_2,y_1)})) \\ \leq\ & \overline{P}(A_{(x_2,y_2)}) - \underline{P}(A_{(x_1,y_2)}) - \overline{P}(A_{(x_2,y_1)}) + \overline{P}(A_{(x_1,y_2)} \cap A_{(x_2,y_1)}). \\ & P(A_{(x_2,y_2)}) - P(A_{(x_1,y_2)}) - \underline{P}(A_{(x_2,y_1)}) + P(A_{(x_1,y_2)} \cap A_{(x_2,y_1)}). \end{aligned}$$

In terms of $\underline{F},\overline{F}$, this means that:

$$\overline{F}(x_2,y_2) + \overline{F}(x_1,y_1) - \overline{F}(x_1,y_2) - \underline{F}(x_2,y_1) \geq 0.$$
$$F(x_2,y_2) + F(x_1,y_1) - \underline{F}(x_1,y_2) - F(x_2,y_1) \geq 0.$$

None of these conditions is sufficient for coherence, as we can see in the following examples.

Example 4.99 *Let us show an example where both $\underline{F}$ and $\overline{F}$ satisfy (I-RI1), (I-RI2) and (I-RI4), but not (I-RI3), and the lower prevision $\underline{P}$ is not coherent. For this aim consider three real numbers $a < b < c$ and the functions $\underline{F}$ and $\overline{F}$ defined by:*

	(a,a)	(a,b)	(a,c)	(b,a)	(b,b)	(b,c)	(c,a)	(c,b)	(c,c)
$\underline{F}$	0	0.3	0.45	0.3	0.6	0.75	0.45	0.8	1
$\overline{F}$	0	0.3	0.5	0.3	0.6	0.85	0.5	0.85	1

Both $\underline{F}$ and $\overline{F}$ are standardized maps. In addition, $\underline{F}$ is a distribution function, and consequently $\underline{F}$ and $\overline{F}$ satisfy (I-RI1) and (I-RI2). It can be checked that (I-RI4) is also satisfied. Assume that their lower probability $\underline{P}_{(\underline{F},\overline{F})}$ is coherent. Then, by Proposition 4.97 *there must be a distribution function* F *between* $\underline{F}, \overline{F}$ *such that* $F(b,c) = \overline{F}(b,c) = 0.85$. *However, this implies that*

$$F(c,c) + F(b,b) - F(b,c) - F(c,b) = 1 + 0.6 - 0.85 - F(c,b) \geq 0 \Rightarrow F(c,b) \leq 0.75.$$

But on the other hand we must have $F(c,b) \geq \underline{F}(c,b) = 0.8$*: this is a contradiction.*

Similarly, if we define $\underline{F}$ *and* $\overline{F}$ *by* $\underline{F}(x,y) = \underline{F}(y,x)$ *and* $\overline{F}(x,y) = \overline{F}(y,x)$, *we obtain an example where (I-RI1), (I-RI2) and (I-RI3) are satisfied but the p-box* $(\underline{F}, \overline{F})$ *is not coherent.*

Example 4.100 *Let us give next an example where* $\underline{F}$ *and* $\overline{F}$ *satisfy conditions (I-RI2) and (I-RI3) and (I-RI4), but not (I-RI1), and the bivariate p-box* $(\underline{F}, \overline{F})$ *is not coherent. For this aim consider three real numbers* $a < b < c$ *and the functions* $\underline{F}$ *and* $\overline{F}$ *defined by:*

	(a,a)	(a,b)	(a,c)	(b,a)	(b,b)	(b,c)	(c,a)	(c,b)	(c,c)
$\underline{F}$	0	0.3	0.4	0.3	0.6	0.6	0.5	0.8	1
$\overline{F}$	0	0.3	0.4	0.3	0.6	0.7	0.5	0.8	1

Both $\underline{F}$ *and* $\overline{F}$ *are standardized functions. They also satisfy conditions (I-RI2) and, since* $\overline{F}$ *is a cumulative distribution function, also conditions (I-RI3) and (I-RI4). Assume that* $(\underline{F}, \overline{F})$ *is coherent. Then, there must be a distribution function* F *such that* $F(b,c) = \underline{F}(b,c) = 0.6$. *Then:*

$$F(b,c) + F(a,b) - F(b,b) - F(a,c) = 0.6 + 0.3 - 0.6 - 0.4 = -0.1 < 0,$$

a contradiction.

Example 4.101 *Finally, let us give an example where* $\underline{F}$ *and* $\overline{F}$ *satisfy (I-RI1) and (I-RI3) and (I-RI4), but not condition (I-RI2), and the bivariate p-box* $(\underline{F}, \overline{F})$ *is not coherent. As in the previous examples, consider three real numbers* $a < b < c$ *and the functions* $\underline{F}$ *and* $\overline{F}$ *defined by:*

	(a,a)	(a,b)	(a,c)	(b,a)	(b,b)	(b,c)	(c,a)	(c,b)	(c,c)
$\underline{F}$	0	0.3	0.4	0.3	0.5	0.7	0.5	0.8	1
$\overline{F}$	0.1	0.3	0.4	0.3	0.5	0.7	0.5	0.8	1

These functions can be easily proven to satisfy (I-RI1), (I-RI3) and (I-RI4). However, they do not satisfy (I-RI2) since:

$$\overline{F}(b,b) + \underline{F}(a,a) - \underline{F}(a,b) - \underline{F}(b,a) = 0.5 + 0 - 0.3 - 0.3 = -0.1 < 0.$$

Then, $\mathcal{P}_{(\underline{F},\overline{F})}$ *is not coherent.*

Next we establish the most important result in this section: a characterisation of the coherence of a bivariate p-box in the case when one of the variables is binary.

Proposition 4.102 *Assume that $\Omega_2 = \{y_1, y_2\}$ is a binary space and $\Omega_1 = \{x_1, \dots, x_n\}$ is finite, and let $(\underline{F}, \overline{F})$ be a bivariate p-box on $\Omega_1 \times \Omega_2$.*

1. *If $\underline{F}, \overline{F}$ satisfy (I-RI1) and (I-RI2), then*
$$\underline{F} = \min \{F \text{ distribution function} : \underline{F} \le F \le \overline{F}\}.$$
2. *If $\underline{F}, \overline{F}$ satisfy (I-RI3) and (I-RI4), then*
$$\overline{F} = \max \{F \text{ distribution function} : \underline{F} \le F \le \overline{F}\}.$$
3. *As a consequence, $(\underline{F}, \overline{F})$ is coherent $\quad \underline{F}, \overline{F}$ satisfy conditions (I-RI1) to (I-RI4).*

Proof First of all, let us check that if $\underline{F}$ and $\overline{F}$ satisfy (I-RI2), then there is a cumulative distribution function F_2 such that $\underline{F} \le F_2$ and $F_2(x_i, y_1) = \underline{F}(x_i, y_1)$ for any $i = 1, \dots, n$. For this aim we define the function F_2 by:

$$\begin{aligned}
&F_2(x_i, y_1) = \underline{F}(x_i, y_1) \text{ for } i = 1, \dots, n,\\
&F_2(x_1, y_2) = \underline{F}(x_1, y_2), \text{ and}\\
&F_2(x_i, y_2) = \underline{F}(x_i, y_2) - \min(0, \Delta^{R_{i-1}}_{\underline{F}}), \text{ for } i = 2, \dots, n, \quad \text{where}\\
&\Delta^{R_{i-1}}_{\underline{F}} = \underline{F}(x_i, y_2) + \underline{F}(x_{i-1}, y_1) - \underline{F}(x_i, y_1) - F_2(x_{i-1}, y_2).
\end{aligned}$$

On the one hand, by definition $F_2(x_i, y_1) = \underline{F}(x_i, y_1)$ for $i = 1, \dots, n$. On the other hand, let us prove that $\underline{F} \le F_2 \le \overline{F}$, $F_2(x_n, y_2) = 1$, F_2 is monotone and $\Delta^{R_{i-1}}_{F_2} \ge 0$, where:
$$\Delta^{R_{i-1}}_{F_2} = F_2(x_i, y_2) + F_2(x_{i-1}, y_1) - F_2(x_i, y_1) - F_2(x_{i-1}, y_2),$$
for $i = 2, \dots, n$. In such a case, F_2 would be a distribution function bounded by $\underline{F}$ and $\overline{F}$.

1. $F_2 \ge \underline{F}$:

   It trivially holds since $-\min(0, \Delta^{R_{i-1}}_{\underline{F}}) \ge 0$.

2. $F_2 \le \overline{F}$:

   For either $i=1$ or $j=1$, $F_2(x_i, y_j) = \underline{F}(x_i, y_j) \le \overline{F}(x_i, y_j)$. When $i, j \ge 2$, and $(i, j) = (n, 2)$, it holds that:
$$\overline{F}(x_i, y_2) \ge F_2(x_i, y_2) \qquad \overline{F}(x_i, y_2) - \underline{F}(x_i, y_2) + \min(\Delta^{R_{i-1}}_{\underline{F}}, 0) \ge 0$$
   This is obvious when $\Delta^{R_{i-1}}_{\underline{F}} \ge 0$. Otherwise, we have to prove that
$$\overline{F}(x_i, y_2) - \underline{F}(x_i, y_2) + \Delta^{R_{i-1}}_{\underline{F}} \ge 0.$$

Thisinequalityholdsif and onlyif:

$$
\begin{aligned}
0 \le\ & \overline{F}(x_i,y_2) - \underline{F}(x_i,y_2) + \underline{F}(x_i,y_2) - \underline{F}(x_i,y_1) - F_2(x_{i-1},y_2) + \underline{F}(x_{i-1},y_1) \\
=\ & F(x_i,y_2) - \underline{F}(x_i,y_1) - F_2(x_{i-1},y_2) + \underline{F}(x_{i-1},y_1).
\end{aligned}
$$

Then, we shall prove that

$$
\overline{F}(x_i,y_2) - \underline{F}(x_i,y_1) - F_2(x_k,y_2) + \underline{F}(x_k,y_1) \ge 0 \tag{4.16}
$$

for any $k = 1,\ldots,i-1$ by indu ction on k.

(a) $k=1$: Equation (4.16) becomes:

$$
\overline{F}(x_i,y_2) - \underline{F}(x_i,y_1) - \underline{F}(x_1,y_2) + \underline{F}(x_1,y_1) \ge 0,
$$

andit holdsfor (I-RI2).

(b) Assume thatEquation (4.16)holds for $k-1$. Then, for $k=1$ Equation (4.16) becomes:

$$
\overline{F}(x_i,y_2) - \underline{F}(x_i,y_1) - \underline{F}(x_k,y_2) + \min(\Delta_{\underline{E}}^{R_{k-1}}, 0) + \underline{F}(x_k,y_1) \ge 0,
$$

and this is positive when $\Delta_{\underline{E}}^{R_{k-1}} \ge 0$ by (I-RI2). Otherwise, it becomes:

$$
\begin{aligned}
\overline{F}(x_i,y_2) & - \underline{F}(x_i,y_1) - \underline{F}(x_k,y_2) + \underline{F}(x_k,y_2) - \underline{F}(x_k,y_1) \\
& - \underline{F}_2(x_{k-1},y_2) + \underline{F}(x_{k-1},y_1) + \underline{F}(x_k,y_1) \\
& = F(x_i,y_2) - \underline{F}(x_i,y_1) - F_2(x_{k-1},y_2) + \underline{F}(x_{k-1},y_1) \ge 0,
\end{aligned}
$$

sinceEquation (4.16) holds for $k-1$.

3. $F_2(x_n,y_2) = 1$:

   In fact:

$$
\begin{aligned}
F_2(x_n,y_2) = 1 \quad & \underline{F}(x_n,y_2) - \min(\Delta_{\underline{E}}^{R_{n-1}},0) = 1 - \min(\Delta_{\underline{E}}^{R_{n-1}},0) = 1 \\
& \Delta_{\underline{E}}^{R_{n-1}} \ge 0 \\
& \underline{F}(x_n,y_2) - \underline{F}(x_n,y_1) - F_2(x_{n-1},y_2) + \underline{F}(x_{n-1},y_1) \\
& = F(x_n,y_2) - \underline{F}(x_n,y_1) - F_2(x_{n-1},y_2) + \underline{F}(x_{n-1},y_1) \ge 0,
\end{aligned}
$$

   which follows from the pro of by induction of Equation (4.16) by putting $i=n$ and $k=n-1$.

4. F_2 is monotone:

   (a) On theone hand, $F_2(x_i,y_1) = \underline{F}(x_i,y_1) \le \underline{F}(x_{i+1},y_1) = F_2(x_i,y_1)$ for any $i = 1,\ldots,n-1$.

(b) $F_2(x_i,y_2) \geq F_2(x_{i-1},y_2)$:

$$\begin{aligned} F_2(x_i,y_2) &= \underline{F}(x_i,y_2) - \min(\Delta^{R_{i-1}}_{\underline{F}},0) \\ &= \max(\underline{F}(x_i,y_2) - \Delta^{R_{i-1}}_{\underline{F}}, \underline{F}(x_i,y_2)) \\ &= \max(F_2(x_{i-1},y_2) + \underline{F}(x_i,y_1) - \underline{F}(x_{i-1},y_1), \underline{F}(x_i,y_2)) \\ &\geq F_2(x_{i-1},y_2) + \underline{F}(x_i,y_1) - \underline{F}(x_{i-1},y_1) \geq F_2(x_{i-1},y_2), \end{aligned}$$

by the monotonicity of $\underline{F}$.

(c) $F_2(x_i,y_2) \geq F_2(x_i,y_1) = \underline{F}(x_i,y_1)$ since

$$F_2(x_i,y_2) \geq \underline{F}(x_i,y_2) \geq \underline{F}(x_i,y_1).$$

5. $\Delta^{R_{i-1}}_{F_2} \geq 0$ for $i = 1,\ldots,n$:

It holds that:

$$\begin{aligned} \Delta^{R_{i-1}}_{F_2} &= F_2(x_i,y_2) - \underline{F}(x_i,y_1) - F_2(x_{i-1},y_2) + \underline{F}(x_{i-1},y_1) \\ &= \underline{F}(x_i,y_2) + \max(-\Delta^{R_{i-1}}_{\underline{F}},0) - \underline{F}(x_i,y_1) - F_2(x_{i-1},y_2) + \underline{F}(x_{i-1},y_1) \\ &= \max(-\Delta^{R_{i-1}}_{\underline{F}},0) + \Delta^{R_{i-1}}_{\underline{F}} = \max(0,\Delta^{R_{i-1}}_{\underline{F}}) \geq 0. \end{aligned}$$

Now, consider the function F_1 defined by:

$$\begin{aligned} &F_1(x_i,y_2) = \underline{F}(x_i,y_2) \text{ for } i = 1,\ldots,n, \\ &F_1(x_i,y_1) = \underline{F}(x_i,y_1) - \min(\Delta^{R_i}_{\underline{F}}, 0), \text{ where} \\ &\Delta^{R_i}_{\underline{F}} = \underline{F}(x_{i+1},y_2) - F_1(x_{i+1},y_1) - \underline{F}(x_i,y_2) + \underline{F}(x_i,y_1), \end{aligned}$$

for $i=n-1,\ldots,1$. If $\underline{F}$ and $\overline{F}$ satisfy (I-RI1), with a similar proof as the one for F_2, we can prove that F_1 is a distribution function bounded by $\underline{F}$ and $\overline{F}$ and, by its definition, $F_1(x_i,y_2) = \underline{F}(x_i,y_2)$ for $i = 1,\ldots,n$. Then, taking into account F_1 and F_2, it holds that:

$$\underline{F} = \min\{F \text{ distribution functions } : \underline{F} \leq F \leq \overline{F}\}.$$

Finally, consider the functions F_3 and F_4, defined by:

$$\begin{aligned} &F_3(x_i,y_2) = \overline{F}(x_i,y_2) \text{ for } i = 1,\ldots,n, \\ &F_3(x_1,y_1) = \overline{F}(x_1,y_1), \text{ and} \\ &F_3(x_i,y_1) = \overline{F}(x_i,y_1) + \min(\Delta^{R_{i-1}}_{\overline{F}},0), \text{ where} \\ &\Delta^{R_{i-1}}_{\overline{F}} = \overline{F}(x_i,y_2) + F_3(x_{i-1},y_1) - \overline{F}(x_{i-1},y_2) - \overline{F}(x_i,y_1) \end{aligned}$$

for $i = 2,\ldots,n$, and:

$$\begin{aligned} &F_4(x_i,y_1) = \overline{F}(x_i,y_1) \text{ for } i = 1,\ldots,n, \\ &F_4(x_n,y_2) = \overline{F}(x_n,y_2), \text{ and} \\ &F_4(x_i,y_1) = \overline{F}(x_i,y_1) + \min(\Delta^{R_i}_{\overline{F}}, 0), \text{ where} \\ &\Delta^{R_i}_{\overline{F}} = \overline{F}(x_{i+1},y_2) + \overline{F}(x_i,y_1) - F_4(x_i,y_2) - \overline{F}(x_{i-1},y_1) \end{aligned}$$

for $i = n-1, \dots, 1$. With a similar proof as the one for F_2, we can check that when $\underline{F}$ and $\overline{F}$ satisfy (I-RI3) (respectively (I-RI4)) F_3 (respectively F_4) is a distribution function bounded by $\underline{F}$ and $\overline{F}$ such that $F_3(x_i, y_2) = \overline{F}(x_i, y_2)$ (respectively $F_4(x_i, y_1) = \overline{F}(x_i, y_1)$) for $i = 1, \dots, n$. Then, this implies that when $\underline{F}$ and $\overline{F}$ satisfy conditions (I-RI3) and (I-RI4) it holds that:

$$\overline{F} = \max\{F \text{ distribution functions} : \underline{F} \leq F \leq \overline{F}\}.$$

Putting the functions F_1, F_2, F_3 and F_4 together, we deduce that when $\underline{F}$ and $\overline{F}$ satisfy (I-RI1) to (I-RI4), $(\underline{F}, \overline{F})$ is a coherent bivariate p-box; the converse implication holds by Proposition 4.98. ∎

As a consequence, we deduce that conditions (I-RI1)–(I-RI4) are also equivalent to the coherence of $(\underline{F}, \overline{F})$ when both variables Ω_1, Ω_2 are binary. In fact, we conjecture that conditions (I-RI1)–(I-RI4) are also equivalent to the coherence of $(\underline{F}, \overline{F})$ in the general case.

To conclude this section, we investigate if the third statement in Theorem 2.35 can be used to characterise coherence in the bivariate case. Let $\underline{F}, \overline{F}$ be standardized maps on $\Omega_1 \times \Omega_2$, and let $\underline{P}_{\underline{F}} : \mathcal{K}_1 \to \mathbb{R}$ and $\underline{P}_{\overline{F}} : \mathcal{K}_2 \to \mathbb{R}$ be the lower probabilities associated with them by Equation (4.13).

Proposition 4.103 *Let $(\underline{F}, \overline{F})$ be a bivariate p-box and let $\underline{P}_{\underline{F}}, \underline{P}_{\overline{F}}$ be the lower previsions they induce on $\mathcal{K}_1, \mathcal{K}_2$, respectively. Then:*

(a) $\underline{P}_{\underline{F}}, \underline{P}_{\overline{F}}$ always avoid sure loss.

(b) $\underline{P}_{\underline{F}}$ is coherent $\Leftrightarrow$ $\underline{P}_{(\underline{F},1)}$ is coherent.

(c) $\underline{P}_{\overline{F}}$ is coherent $\Leftrightarrow$ $\underline{P}_{(I_{(x,y)},\overline{F})}$ is coherent.

(d) $\underline{P}_{(\underline{F},\overline{F})}$ coherent $\Rightarrow$ $\underline{P}_{\underline{F}}, \underline{P}_{\overline{F}}$ coherent.

Proof

(a) To see that $\underline{P}_{\underline{F}}$ and $\underline{P}_{\overline{F}}$ always avoid sure loss, it suffices to take into account that the constant map on 1 is a distribution function that dominates $\underline{F}$ and that $I_{(x,y)}$ is a distribution function that is dominated by $\overline{F}$.

(b) The lower probability $\underline{P}_{\underline{F}}$ is coherent if and only if for every $(x, y) \in \Omega_1 \times \Omega_2$ there is a distribution function $F \geq \underline{F}$ such that $F(x, y) = \underline{F}(x, y)$. The condition $F \geq \underline{F}$ is equivalent to $\underline{F} \leq F \leq 1$, and on the other hand the constant map on 1 is trivially a distribution function. We deduce from Proposition 4.97 that $\underline{P}_{(\underline{F},1)}$ is coherent if and only if $\underline{F}$ is the lower envelope of the distribution functions in $(\underline{F}, 1)$ and as a consequence we have the equivalence.

(c) The lower probability $\underline{P}_{\overline{F}}$ is coherent if and only if for every $(x, y) \in \Omega_1 \times \Omega_2$ there is a distribution function $F \leq \overline{F}$ such that $F(x, y) = \overline{F}(x, y)$. The condition $F \leq \overline{F}$ is equivalent to $I_{(x,y)} \leq F \leq \overline{F}$, and on the other hand the map $I_{(x,y)}$ is trivially a distribution function. We deduce from Proposition 4.97 that $\underline{P}_{(I_{(x,y)},\overline{F})}$ is coherent if and only if $\overline{F}$ is the upper envelope of the distribution functions in $(I_{(x,y)}, \overline{F})$, and as a consequence we have the equivalence.

(d) This statement follows from the previous two and from Proposition 4.97, taking into account that the set of distribution functions $(\underline{F}, \overline{F})$ is the intersection of the sets $(\underline{F}, 1)$ and $(I_{(x,y)}, \overline{F})$. ∎

To see that the converse in the fourth statement does not hold, consider the following example.

Example 4.104 *Consider now the functions $\underline{F}$ and $\overline{F}$ of Example 4.100. To see that $(\underline{F}, 1)$ is coherent, it suffices to take into account that $\underline{F}$ is the lower envelope of the distribution functions F_1, F_2 given by:*

	(a, a)	(a, b)	(a, c)	(b, a)	(b, b)	(b, c)	(c, a)	(c, b)	(c, c)
F_1	0	0.3	0.4	0.3	0.6	0.7	0.5	0.8	1
F_2	0.1	0.4	0.4	0.3	0.6	0.6	0.5	0.8	1

while the constant map on 1 is trivially a distribution function.

Similarly, since both $I_{(c,c)}$ and $\overline{F}$ are distribution functions, we deduce that $(I_{(c,c)}, \overline{F})$ is also coherent. However, we saw in Example 4.100 that $(\underline{F}, \overline{F})$ are not coherent.

This shows that one of the equivalences in Theorem 2.35 does not extend to the bivariate case. Moreover, we can see from this example that the coherence of $\underline{P}_{\underline{F}}$ does not imply that $\underline{F}$ is a distribution function: we have that $\underline{F}(a,b) + \underline{F}(b, c) < \underline{F}(a,c) + \underline{F}(b,b)$. In a similar way (using for instance Example 4.99) we can see that the coherence of $\underline{P}_{\overline{F}}$ does not imply that $\overline{F}$ is a distribution function.

Another consequence is that whenever (I-RI1)–(I-RI4) characterise the coherence of $(\underline{F}, \overline{F})$ (as is for instance the case in Proposition 4.102), it holds that $\underline{P}_{\overline{F}}$ is coherent for any a standardized function $\overline{F}$, because they hold trivially whenever $\underline{F}$ is the indicator function $I_{(x,y)}$. On the other hand, $\underline{P}_{\underline{F}}$ may not be coherent: consider $\Omega_1 = \Omega_2 = \{0, 1\}$ and $\underline{F}$ given by:

	$(0, 0)$	$(0, 1)$	$(1, 0)$	$(1, 1)$
$\underline{F}$	0	0.6	0.6	1

Then there is no distribution function $F \geq \underline{F}$ satisfying $F(0, 0) = \underline{F}(0, 0) = 0$, because then

$$F(1,1) + F(0,0) = 1 < 1.2 \leq F(0, 1) + F(1, 0).$$

2-monotonicity In the univariate case, the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ associated with a p-box is completely monotone [198]. As we saw in Definition 2.40, this means, in particular, that for every pair of events A, B in its domain it holds that

$$\underline{P}_{(\underline{F},\overline{F})}(A \cup B) + \underline{P}_{(\underline{F},\overline{F})}(A \cap B) \geq \underline{P}_{(\underline{F},\overline{F})}(A) + \underline{P}_{(\underline{F},\overline{F})}(B),$$

provided also $A \cup B$ and $A \cap B$ belong to the domain. 2-monotone capacities have been studied in detail in [53, 204], among others. They satisfy the property of *comonotone additivity*, which is of interest in economy ([35, 203]).

In the univariate case, we can assume without loss of generality that the domain of the lower probability induced by the p-box is a lattice (see [198] for more details), and this allows us to apply the results from [53]. This is not the case for bivariate p-boxes: the domain $\mathcal{K}$ of $\underline{P}_{(\underline{F},\overline{F})}$ is not a lattice, so if we want to use the results in [53] we need to take the natural extension of $\underline{P}_{(\underline{F},\overline{F})}$. By the Envelope Theorem (Theorem 2.30) and Proposition 4.97, this natural extension is the lower envelope of the set

$$\{P_F : F \text{ distribution function}, \underline{F} \leq F \leq \overline{F}\},$$

where P_F is the finitely additive probability associated with the distribution function F by means of Equation (4.15).

However, and as the following example shows, in the bivariate case it could be that the lower probability associated with the p-box $(\underline{F}, \overline{F})$ is coherent but not 2-monotone, even if both $\underline{F}, \overline{F}$ are distribution functions:

Example 4.105 *Consider $\Omega_1 = \Omega_2 = \{0, 1\}$, and let $\underline{F}, \overline{F}: \Omega_1 \times \Omega_2 \to [0, 1]$ be the standardized maps given by:*

	(0, 0)	(0, 1)	(1, 0)	(1, 1)
$\underline{F}$	0	0	0.5	1
$\overline{F}$	0.25	0.25	0.5	1

Then, both $\underline{F}, \overline{F}$ are distribution functions, because

$$\underline{F}(1, 1) + \underline{F}(0, 0) - \underline{F}(0, 1) - \underline{F}(1, 0) = 0;$$
$$\overline{F}(1, 1) + \overline{F}(0, 0) - \overline{F}(0, 1) - \overline{F}(1, 0) = 0.25 > 0;$$

and the other comparisons are trivial.

Now, in the particular case of binary spaces the correspondence between distribution functions and finitely additive probabilities in Equation (4.15) means that any distribution function F on $\Omega_1 \times \Omega_2$ determines uniquely a probability mass function on $\mathcal{P}(\Omega_1) \times \mathcal{P}(\Omega_2)$ by:

$$P_F(\{(0,0)\}) = F(0, 0).$$
$$P_F(\{(0,1)\}) = F(0, 1) - F(0, 0).$$
$$P_F(\{(1,0)\}) = F(1, 0) - F(0, 0).$$
$$P_F(\{(1,1)\}) = 1 - P_F(\{(0, 1)\}) - P_F(\{(1, 0)\}) - P_F(\{(0, 0)\})$$
$$= F(1, 1) - F(0, 1) - F(1, 0) + F(0, 0) \geq 0.$$

Let $\mathcal{F}$ be the set of distribution functions that lie between $\underline{F}$ and $\overline{F}$, and let us define

$$\mathcal{M}_{\mathcal{F}} := \{P_F : F \in \mathcal{F}\} .$$

Then $\underline{P}_{(\underline{F},\overline{F})}$ is the lower envelope of $\mathcal{M}_{\mathcal{F}}$ on $\mathcal{K}$ and so is its natural extension $\underline{E}$. Let us show that $\underline{E}$ is not 2-monotone.

Since $\overline{F}(1,0) = 0.5$, $\overline{F}(0,1) = 0.25$ and $\underline{F}(1,1) = 1$, any map F bounded between $\underline{F}$ and $\overline{F}$ will satisfy $F(1,0) + F(0,1) \leq F(0,0) + F(1,1)$, so it will be a distribution function as soon as it is monotone. In other words, $\mathcal{F} = \{F \text{ monotone} : \underline{F} \leq F \leq \overline{F}\}$.

Denote $a = \{(0,0)\}$, $b = \{(0,1)\}$, $c = \{(1,0)\}$, $d = \{(1,1)\}$ and take $A = \{a,c\}$ and $B = \{c,d\}$. Any monotone map F bounded by $\underline{F},\overline{F}$ induces the mass function $(P(a), P(b), P(c), P(d))$, where:

$$P(a) \in [0, 0.25], \qquad P(a) + P(b) \in [0, 0.25],$$
$$P(a) + P(c) = 0.5, \qquad P(a) + P(b) + P(c) + P(d) = 1.$$

Then:

$$\begin{aligned}\mathcal{M}_{\mathcal{F}} &= \{(P_F(a), P_F(b), P_F(c), P_F(d)) : F \in (\underline{F},\overline{F})\} \\ &= \{(\lambda, \nu - \lambda, 0.5 - \lambda, 0.5 - \nu + \lambda) : \nu \in [0, 0.25], \lambda \in [0, \nu]\},\end{aligned}$$

and as a consequence:

- $\underline{E}(A) = \underline{E}(\{a,c\}) = 0.5$.
- $\underline{E}(B) = \min\{P(c) + P(d) : P \in \mathcal{M}_{\mathcal{F}}\} = 0.75$, *considering the mass function* $P = (0.25, 0, 0.25, 0.5)$
- $\underline{E}(A \cup B) = \min\{P(a) + P(c) + P(d) : P \in \mathcal{M}_{\mathcal{F}}\} = 0.75$, *with* $P = (0, 0.25, 0.5, 0.25)$
- $\underline{E}(A \cap B) = \min\{P(c) : P \in \mathcal{M}_{\mathcal{F}}\} = 0.25$, *considering the mass function* $P = (0.25, 0, 0.25, 0.5)$

This means that $\underline{E}(A \cup B) + \underline{E}(A \cap B) < \underline{E}(A) + \underline{E}(B)$ and therefore the lower probability induced by the p-box $(\underline{F},\overline{F})$ is not 2-monotone.

Interestingly, in this example the lower probability $\underline{E}$ does not coincide with the lower envelope of $\min\{P_{\underline{F}}, P_{\overline{F}}\}$: these are associated with the mass functions $P_{\underline{F}} = (0, 0, 0.5, 0.5)$ and $P_{\overline{F}} = (0.25, 0, 0.25, 0.5)$, so

$$\min\{P_{\underline{F}}(A \cup B), P_{\overline{F}}(A \cup B)\} = 1 > 0.75 = \underline{E}(A \cup B).$$

This means that even if the p-box is determined by the distribution functions $\underline{F},\overline{F}$, the same does not apply to its associated lower probability.

On the other hand, when the bivariate p-box determines a 2-monotone lower probability, it is not too difficult to show that $\underline{F}$ is indeed a distribution function. Note here the difference with the case where we only require that the lower probability is coherent, discussed in Section 4.3.1.

Proposition 4.106 ([185, Lemma 6]) *Assume that the natural extension of the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ induced by the bivariate p-box $(\underline{F},\overline{F})$ by Equation (4.14) is 2-monotone. Then $\underline{F}$ is a distribution function.*

However, the standardized map $\overline{F}$ of the p-box determined by a 2-monotone lower probability is not necessarily a distribution function.

Example 4.107 *Consider the upper probability defined by* $\overline{P}(A) = \min((1+\delta)P(A), 1)$ *for every* $A \in \mathcal{P}(\Omega_1 \times \Omega_2)$, *where* $\delta > 0$,

$$K \subseteq \{A_{(x,y)} : x \in \Omega_1, y \in \Omega_2\},$$

and P is a probability measure. This corresponds to Pari-mutuel model (see [205, Section 2.9.3]) and it is known that $\overline{P}$ is 2-alternating. Consider the random variables X and Y defined on $\Omega_1 = \Omega_2 = \{a, b, c\}$, where $a < b < c$, probability P and value of $\delta = 0.25$:

$X \setminus Y$	a	b	c
a	0.1	0	0.15
b	0.2	0.2	0.05
c	0.15	0.1	0.05

Joint probability distribution

$X \setminus Y$	a	b	c
a	0.1	0.1	0.25
b	0.3	0.5	0.7
c	0.45	0.75	1

Joint distribution function

In this situation, $\overline{F}$ is not a precise cumulative distribution function:

$$\overline{F}(3,3) + \overline{F}(2,2) - \overline{F}(3,2) - \overline{F}(2,3) = 1 + 0.625 - 0.9375 - 0.875 < 0.$$

Remark 4.108 *One interesting case is that when the bivariate p-box is precise, that is, when the standardized maps $\underline{F}, \overline{F}$ coincide. In that case, we obviously have that $(\underline{F}, \overline{F})$ avoids sure loss if and only if it is coherent, and if and only if $\underline{F} = \overline{F}$ is a bivariate distribution function. When Ω_1 and Ω_2 are finite, it follows from Equation (4.15) that this distribution function has a unique extension to the power set of $\Omega_1 \times \Omega_2$; this means that in that case the lower probability associated with $(\underline{F}, \overline{F})$ is linear.*

Note however, that a distribution function does not determine uniquely its associated finitely additive probability, not even in the univariate case; this is a problem that has been explored in detail in [133].

4.3.2 Imprecise copulas

One particular case where bivariate p-boxes can arise is in the combination of two marginal p-boxes. In this section, we shall explore this case in detail, by studying the properties of a number of bivariate p-boxes with given marginals: the most conservative one, that shall be obtained by means of the Fréchet bounds and the notion of natural extension, and also the one corresponding to model notion of independence. In both cases, we shall see that the bivariate model can be derived by means of an appropriate extension of the notion of copula.

Related results can be found in [198, Section 7], with one fundamental difference: in [198], the authors assume the existence of a total preorder on the product space $\Omega_1 \times \Omega_2$ that is compatible with the orders in Ω_1, Ω_2; while here we shall only consider the partial order given by

$$(x_1, y_1) \leq (x_2, y_2) \quad x_1 \leq x_2 \text{ and } y_1 \leq y_2.$$

An imprecise version of Sklar's theorem

Taking into account our previous results, we see that the combination of the marginal p-boxes into a bivariate one is related to the combination of marginal lower probabilities into a joint one. This is a problem that has been studied in detail under some conditions of independence [52].

Remember that Sklar's Theorem (see Theorem 2.27) stated that given two random variables X and Y with associated cumulative distribution functions F_X and F_Y, there exists a copula C such that the joint distribution function, named F, can be expressed by:

$$F(x, y) = C(F_X(x), F_Y(y)) \text{ for any } x, y.$$

Moreover, the copula is unique on $Rang(F_X) \times Rang(F_Y)$. Conversely, any transformation of marginal distribution functions by means of a copula produces a bivariate distribution function.

Next, we introduce the notion of imprecise copula. It is a simple generalisation of precise copulas; the main difference lies in the rectangle inequality that has been replaced by its four imprecise extensions of (I-RI1)–(I-RI4).

Definition 4.109 *A pair of functions* $\underline{C}, \overline{C} : [0, 1] \times [0, 1] \to [0, 1]$ *is called an* imprecise copula *if:*

- *Both* $\underline{C}$ *and* $\overline{C}$ *are component-wise increasing.*
- $\underline{C} \leq \overline{C}$.

- $\underline{C}(0,u) = \overline{C}(0,u) = 0 = \overline{C}(v,0) = \underline{C}(v,0)\ \forall u, v \in [0,1]$
- $\underline{C}(1,u) = \overline{C}(1,u) = u$ and $\overline{C}(v,1) = \underline{C}(v,1) = v\ \forall u \in S_2, v \in S_1$.
- $\underline{C}$ *and* $\overline{C}$ *satisfy the following conditions for any* $x_1, x_2, y_1, y_2 \in [0,1]$ *such that* $x_1 \le x_2$ *and* $y_1 \le y_2$:

$$
\begin{aligned}
&(\mathbf{I-CRI1}): && \overline{C}(x_1,y_1) + \underline{C}(x_2,y_2) \ge \underline{C}(x_1,y_2) + \underline{C}(x_2,y_1).\\
&(\mathbf{I-CRI2}): && \underline{C}(x_1,y_1) + \overline{C}(x_2,y_2) \ge \underline{C}(x_1,y_2) + \underline{C}(x_2,y_1).\\
&(\mathbf{I-CRI3}): && \overline{C}(x_1,y_1) + \overline{C}(x_2,y_2) \ge \underline{C}(x_1,y_2) + \overline{C}(x_2,y_1).\\
&(\mathbf{I-CRI4}): && \overline{C}(x_1,y_1) + \overline{C}(x_2,y_2) \ge \overline{C}(x_1,y_2) + \underline{C}(x_2,y_1).
\end{aligned}
$$

$\underline{C}$ *and* $\overline{C}$ *shall be named the* lower *and the* upper copulas*, respectively.*

Note that monotonicity and condition $\underline{C} \le \overline{C}$ may not be imposed in the definition of imprecise copula: on the one hand, $\underline{C} \le \overline{C}$ can be derived from conditions (I-CRI1) to (I-CRI4): for any $x, y \in [0,1]$ (I-CRI1) assures that

$$\overline{C}(x,y) + \underline{C}(x,y) \ge \underline{C}(x,y) + \underline{C}(x,y),$$

that is equivalent to $\overline{C}(x,y) \ge \underline{C}(x,y)$. Furthermore, taking $0 \le x$ and $y_1 \le y_2$ and applying (I-CRI1) we obtain that $\underline{C}$ is increasing in the second component. Similarly, using conditions (I-CRI1) to (I-CRI4) we obtain that both $\underline{C}$ and $\overline{C}$ are increasing in each component.

As next result shows, one way of obtaining imprecise copulas is by taking the infimum and supremum of sets of copulas, or just simply by considering two ordered copulas.

Proposition 4.110 *Let* $\mathcal{C}$ *be a non-empty set of copulas. Take* $\underline{C}$ *and* $\overline{C}$ *defined by:*

$$\underline{C}(x,y) = \inf_{C \in \mathcal{C}} C(x,y) \text{ and } \overline{C}(x,y) = \sup_{C \in \mathcal{C}} C(x,y)$$

for any (x,y). *Then,* $(\underline{C}, \overline{C})$ *forms an imprecise copula. Moreover, if* C_1 *and* C_2 *are two copulas such that* $C_1 \le C_2$, *then* (C_1, C_2) *also forms an imprecise copula.*

Proof Consider $\mathcal{C}$ a non-empty set of copulas, and let $\underline{C}$ and $\overline{C}$ denote their infimum and supremum. Since any copula is in particular a bivariate cumulative distribution function, $(\underline{C}, \overline{C})$ forms a bivariate p-box. Hence, $\underline{C}$ and $\overline{C}$ satisfy $\underline{C} \le \overline{C}$, monotonicity, the boundary conditions and (I-CRI1) to (I-CRI4).

In particular, if we consider two copulas C_1 and C_2 such that $C_1 \le C_2$, the previous result applies, being C_1 and C_2 the infimum and supremum, respectively. ∎

Let us see to which extent Sklar's theorem also holds in an imprecise framework. For this aim, we start by considering marginal imprecise distributions, described by (univariate) p-boxes, and we use imprecise copulas to obtain a bivariate p-box that generates a coherent lower probability.

Proposition 4.111 *Let $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$ be two marginal p-boxes on respective spaces Ω_1, Ω_2, and let $\mathcal{C}$ be a set of copulas. Define the bivariate p-box $(\underline{F}, \overline{F})$ by:*

$$\underline{F}(x, y) = \inf_{C \in \mathcal{C}} C(\underline{F}_X(x), \underline{F}_Y(y)) \text{ and } \overline{F}(x, y) = \sup_{C \in \mathcal{C}} C(\overline{F}_X(x), \overline{F}_Y(y)) \tag{4.17}$$

for any (x, y), and let $\underline{P}$ be its associated lower probability by Equation (4.14). Then, $\underline{P}$ is a coherent lower probability. Moreover,

$$\underline{F}(x, y) = \underline{C}(\underline{F}_X(x), \underline{F}_Y(y)) \text{ and } \overline{F}(x, y) = \overline{C}(\overline{F}_X(x), \overline{F}_Y(y)),$$

where $\underline{C}(x, y) = \inf_{C \in \mathcal{C}} C(x, y)$ and $\overline{C}(x, y) = \sup_{C \in \mathcal{C}} C(x, y)$.

Proof Given $C \in \mathcal{C}$, $F_1 \in (\underline{F}_X, \overline{F}_X)$ and $F_2 \in (\underline{F}_Y, \overline{F}_Y)$, the bivariate distribution function $C(F_1, F_2)$ is bounded by $\underline{F}, \overline{F}$. Applying Proposition 4.93, we deduce that $\underline{P}$ avoids sure loss. Let us now check that it is also coherent. Fix (x, y) in $\Omega_1 \times \Omega_2$. Since the marginal p-boxes $(\underline{F}_X, \overline{F}_X)$, $(\underline{F}_Y, \overline{F}_Y)$ are coherent, there are $F_1 \in (\underline{F}_X, \overline{F}_X)$ and $F_2 \in (\underline{F}_Y, \overline{F}_Y)$ such that $F_1(x) = \underline{F}_X(x)$ and $F_2(y) = \underline{F}_Y(y)$. As a consequence,

$$\underline{F}(x, y) = \inf_{C \in \mathcal{C}} C(\underline{F}_X(x), \underline{F}_Y(y)) = \inf_{C \in \mathcal{C}} C(F_1(x), F_2(y)),$$

and since $C(F_1, F_2) \in (\underline{F}, \overline{F})$ for every $C \in \mathcal{C}$, it then follows from monotonicity that $\underline{F}$ is the lower envelope of the set $\{F \text{ distribution function} : \underline{F} \leq F \leq \overline{F}\}$. Similarly, we can also prove that

$$\overline{F} = \sup \{F \text{ distribution function} : \underline{F} \leq F \leq \overline{F}\}.$$

Applying now Proposition 4.97, we deduce that $\underline{P}$ is coherent. ∎

In particular, when the information about the marginal distribution is precise, and it is given by the distribution functions F_X and F_Y, the bivariate p-box in the above proposition is given by

$$\underline{F}(x, y) = \inf_{C \in \mathcal{C}} C(F_X(x), F_Y(y)) \text{ and } \overline{F}(x, y) = \sup_{C \in \mathcal{C}} C(F_X(x), F_Y(y))$$

for any $(x, y) \in \Omega_1 \times \Omega_2$.

Remark 4.112 *This result generalises [167, Theorem 2.4], where the authors only focused on the functions $\underline{F}$ and $\overline{F}$, showing that*

$$\underline{F}(x, y) = \underline{C}(F_X(x), F_Y(y)) \text{ and } \overline{F}(x, y) = \overline{C}(F_X(x), F_Y(y)).$$

Proposition 4.111 establishes moreover the coherence of the joint lower probability, and it is more general than [167, Theorem 2.4] since we are assuming the existence of imprecision in the marginal distribution, that we model by means of p-boxes.

Using these results, we can give the form of the credal set $\mathcal{M}(\underline{P})$ (that is, the set of dominating probabilities) associated with the lower probability $\underline{P}$. Note that, in the sequel, we can assume that the probabilities in $\mathcal{M}(\underline{P})$ are defined on a suitable set of events, larger than the domain of $\underline{P}$. Hence, the domains of $\underline{P}$ and of the probabilities in $\mathcal{M}(\underline{P})$ do not necessarily coincide.

Corollary 4.113 *Under the assumptions of Proposition 4.111, the credal set $\mathcal{M}(\underline{P})$ of the lower probability $\underline{P}$ is given by:*

$$\{P \text{ probability} \mid \underline{C}(\underline{F}_X(x),\underline{F}_Y(y)) \leq F_P(x,y) \leq \overline{C}(\overline{F}_X(x),\overline{F}_Y(y)) \;\forall x,y\}.$$

Proof By Proposition 4.97, we know that $\underline{P}$ is coherent if and only if $\underline{F}$ and $\overline{F}$ are the lower and the upper envelopes of the set

$$\{F \text{ distribution function} \mid \underline{F} \leq F \leq \overline{F}\}.$$

From this, the thesis follows simply by replacing the lower and upper distribution functions by their expressions in terms of $\underline{C}$ and $\overline{C}$. ■

Next, we investigate whether the second part of Sklar's theorem also holds, meaning whether any bivariate p-box can be obtained as the combination of its marginals by means of an imprecise copula. A partial result in this sense has been established in [185, Theorem 9]. The next example shows that this result cannot be generalised to arbitrary p-boxes.

Example 4.114 *Consider $\Omega_1 = \{x_1,x_2,x_3\}$, $\Omega_2 = \{y_1,y_2\}$ with $x_1 < x_2 < x_3$, $y_1 < y_2$ and let P_1,P_2 be the probability measures associated with the mass functions:*

	(x_1,y_1)	(x_2,y_1)	(x_1,y_2)	(x_2,y_2)	(x_3,y_1)	(x_3,y_2)
P_1	0.2	0	0.3	0	0	0.5
P_2	0.1	0.2	0.5	0.1	0	0.1

Let $\underline{P} = \min\{P_1,P_2\}$. Then its associated p-box satisfies $\underline{F}(x_1) = \underline{F}(x_2) = 0.5$ and $\underline{F}(y_1) = 0.2$ while $\underline{F}(x_1,y_1) = 0.1 < \underline{F}(x_2,y_1) = 0.2$. Hence, there is no function C such that $\underline{F}(x_1,y_1) = C(\underline{F}(x_1),\underline{F}(y_1)) = C(\underline{F}(x_2),\underline{F}(y_1)) = \underline{F}(x_2,y_1)$. Consequently, the lower distribution in the bivariate p-box cannot be expressed as a function of its marginals.

Obviously, when both $\underline{F}, \overline{F}$ are bivariate distribution functions , we can express them as a function of their marginals b ecause of Sklar's theorem; the exam ple shows that this is no longer possible when they are simply standardized functions.

Nexttheorem summarises the results ofthisparagraph.

Theorem 4.115(Imprecise version of Sklar's Theorem) *Consider a set of copulas $\mathcal{C}$ and twomarginal p-boxes $(\underline{F}_X, \overline{F}_X)$. The functions $\underline{F}$ and $\overline{F}$ defined by*

$$\underline{F}(x, y) = \inf_{C \in \mathcal{C}} C(\underline{F}_X(x), \underline{F}_Y(y)) \text{ and}$$
$$\overline{F}(x, y) = \sup_{C \in \mathcal{C}} C(\overline{F}_X(x), \overline{F}_Y(y))$$

form abivariate p-box whose marginals are $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$. Furthermore, the lower probability associated wit h this bivariate p-box is coherent.

However, given a bivariatep-box $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$, theremay not be an imprecise copula $(\underline{C}, \overline{C})$ that generates $(\underline{F}, \overline{F})$ fromitsmarginals, evenwhenits associated lower probability is coherent.

Natural extension and indep endent pro ducts

In this section we consider two particular combinations of the marginal p-b oxes into the bivariate one. First of all, we consider the case where there is no information ab out the copula that links themarginal distribution functions.

Lemma 4.116 *Considerthe univariate p-boxes $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$, and let $\underline{P}$ be the lowerprevision defined on*

$$\mathcal{A} = \{A_{(x,y)}, A^c_{(x,y)}, A_{(x,y)}, A^c_{(x,y)} : x, y \in \mathbb{R}\} \tag{4.18}$$

by

$$\underline{P}(A_{(x,y)}) = \underline{F}_X(x) \qquad \underline{P}(A^c_{(x,y)}) = 1 - \overline{F}_X(x). \tag{4.19}$$
$$\underline{P}(A_{(x,y)}) = \underline{F}_Y(y) \qquad \underline{P}(A^c_{(x,y)}) = 1 - \overline{F}_Y(y).$$

Then:

1. *$\underline{P}$ is a coherent lowerprobability.*
2. *$\mathcal{M}(\underline{P}) = \mathcal{M}(C_L, C_M)$, where $\mathcal{M}(C_L, C_M)$ is given by*

$$\{P \text{ prob.} \mid F_P(x, y) \in [C_L(\underline{F}_X(x), \underline{F}_Y(y)), C_M(\overline{F}_X(x), \overline{F}_Y(y))]\}.$$

Proof Let C_P denote the product copula, and let $\underline{P}_{C_P}$ be the coherent lower probability on $\mathcal{K}$ that results from Proposition 4.111, taking $\mathcal{C} = \{C_P\}$. Then $\underline{P}$ coincides with $\underline{P}_{C_P}$ in $\mathcal{A}$, and consequently $\underline{P}$ is coherent.

On the other hand, let us check the equality between the credal sets $\mathcal{M}(\underline{P})$ and $\mathcal{M}(C_L, C_M)$ (note that both sets are trivially non-empty).

- Let P be a probability in $\mathcal{M}(C_L, C_M)$, and let F_P be its associated distribution function. Then it holds that:

$$F_P(x, y^*) \in [C_L(\underline{F}_X(x), 1), C_M(\overline{F}_X(x), 1)] = [\underline{F}_X(x), \overline{F}_X(x)].$$
$$F_P(x^*, y) \in [C_L(1, \underline{F}_Y(y)), C_M(1, \overline{F}_Y(y))] = [\underline{F}_Y(y), \overline{F}_Y(y)].$$

  Thus, the marginal distribution functions of F_P belong to the p-boxes $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$. As a consequence, $P \in \mathcal{M}(\underline{P})$.

- Conversely, let P be a probability on $\mathcal{M}(\underline{P})$, and let F_P be its associated distribution function. Then, Sklar's Theorem assures that there is a (precise) copula C such that $F_P(x, y) = C(F_P(x, y^*), F_P(x^*, y))$ for every $(x, y) \in \Omega_1 \times \Omega_2$. Hence,

$$\begin{aligned} C_L(\underline{F}_X(x), \underline{F}_Y(y)) &\le C_L(F_P(x, y^*), F_P(x^*, y)) \le C(F_P(x, y^*), F_P(x^*, y)) \\ &\le C(\overline{F}_X(x), \overline{F}_Y(y)) \le C_M(\overline{F}_X(x), \overline{F}_Y(y)), \end{aligned}$$

  taking into account that any copula lies between C_L and C_M. We conclude that $P \in \mathcal{M}(C_L, C_M)$ and as a consequence both sets coincide. ■

From this result we can immediately derive the expression of the *natural extension* [205] of two marginal p-boxes, that is the least-committal (i.e., the most imprecise) coherent lower probability that extends $\underline{P}$ to a larger domain.

Proposition 4.112 *Let $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$ be two univariate p-boxes. Let $\underline{P}$ be the lower prevision defined on the set $\mathcal{A}$ given by Equation* (4.18) *by means of Equation* (4.19). *Then, the natural extension $\underline{E}$ of $\underline{P}$ to $\mathcal{K}$ is given by*

$$\underline{E}(A_{(x,y)}) = C_L(\underline{F}_X(x), \underline{F}_Y(y)) \text{ and } \underline{E}(A^c_{(x,y)}) = 1 - C_M(\overline{F}_X(x), \overline{F}_Y(y)).$$

The bivariate p-box $(\underline{F}, \overline{F})$ associated with $\underline{E}$ is given by:

$$\underline{F}(x, y) = C_L(\underline{F}_X(x), \underline{F}_Y(y)) \text{ and } \overline{F}(x, y) = C_M(\overline{F}_X(x), \overline{F}_Y(y)).$$

Proof On the one hand, the lower prevision $\underline{P}$ is coherent from the previous lemma, and in addition its associated credal set is $\mathcal{M}(\underline{P}) = \mathcal{M}(C_L, C_M)$. The natural extension of $\underline{P}$ to the set $\mathcal{K}$ is given by:

$$\begin{aligned} \underline{E}(A_{(x,y)}) &= \inf_{P \in \mathcal{M}(\underline{P})} F_P(x, y) = \inf_{P \in \mathcal{M}(C_L, C_M)} F_P(x, y) = C_L(\underline{F}_X(x), \underline{F}_Y(y)). \\ \underline{E}(A^c_{(x,y)}) &= \inf_{P \in \mathcal{M}(\underline{P})} (1 - P(A_{(x,y)})) = 1 - \sup_{P \in \mathcal{M}(\underline{P})} F_P(x, y) \\ &= 1 - \sup_{P \in \mathcal{M}(C_L, C_M)} F_P(x, y) = 1 - C_M(\overline{F}_X(x), \overline{F}_Y(y)). \end{aligned}$$

The second part is an immediate consequence of the first. ∎

Recall that Proposition 4.110 assures that every pair of copulas C_1 and C_2 satisfying $C_1 \leq C_2$ (in particular C_L and C_M) forms an imprecise copula (C_1, C_2).

Until now, we have studied how to build the joint p-box $(\underline{F}, \overline{F})$ from two given marginals $(\underline{F}_X, \overline{F}_X)$, $(\underline{F}_Y, \overline{F}_Y)$, when we have no information about the interaction between the underlying variables X and Y: we have argued that we should use in that case the natural extension of the associated coherent lower probabilities, which corresponds to combining the compatible univariate distribution functions by means of all the possible copulas, and then considering the lower envelope.

Next, we consider another case of interest: that where the variables X and Y are assumed to be independent. Consider marginal p-boxes $(\underline{F}_X, \overline{F}_X)$, $(\underline{F}_Y, \overline{F}_Y)$, and let $\underline{P}_X, \underline{P}_Y$ the coherent lower probabilities they induce by means of Equation (2.17). We shall also use this notation to refer to their natural extensions, so that

$$\underline{P}_X := \min\{P : P(A_x) \in [\underline{F}_X(x), \overline{F}_X(x)] \ \forall x \in \Omega_1\} \text{ and}$$
$$\underline{P}_Y := \min\{P : P(A_y) \in [\underline{F}_Y(y), \overline{F}_Y(y)] \ \forall y \in \Omega_2\}.$$

Under imprecise information, there is more than one way to model the notion of independence; see [47] for a survey on this topic. Because of this, there is more than one manner in which we can say that a coherent lower prevision $\underline{P}$ on the product space is an independent product of its marginals $\underline{P}_X, \underline{P}_Y$. This was studied in some detail in [52]. In the remainder of this paragraph, we shall follow that paper into assuming that the spaces Ω and Ω are finite. We recall thus the following definitions.

Definition 4.118 *Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}(\Omega_1 \times \Omega_2)$ with marginals $\underline{P}_X, \underline{P}_Y$. We say that $\underline{P}$ is an* independent product *when it is coherent with the conditional lower previsions $\underline{P}_X(\cdot|\Omega_2), \underline{P}_Y(\cdot|\Omega_1)$ derived from $\underline{P}_X, \underline{P}_Y$ by* epistemic irrelevance*, meaning that*

$$\underline{P}_X(f|y) := \underline{P}_X(f(\cdot, y)) \text{ and } \underline{P}_Y(f|x) := \underline{P}_Y(f(x, \cdot)) \ \forall f \in \mathcal{L}(\Omega_1 \times \Omega_2), x \in \Omega_1, y \in \Omega_2.$$

One example of independent product is the *strong product*, given by

$$\underline{P}_X \boxtimes \underline{P}_Y := \inf\{P_X \times P_Y : P_X \geq \underline{P}_X, P_Y \geq \underline{P}_Y\}.$$

This is the joint model satisfying the notion of *strong independence*. However, it is not the only independent product, nor is it the smallest one. In fact, the smallest independent product of the marginal coherent lower previsions $\underline{P}_X, \underline{P}_Y$ is their *independent natural extension*, which is given by

$$\begin{aligned}(\underline{P}_X \otimes \underline{P}_Y)(f) \\ = \sup\{\mu : f - \mu \geq g - \underline{P}_X(g|\Omega_2) + h - \underline{P}_Y(h|\Omega_1) \text{ for some } g, h \in \mathcal{L}(\Omega_1 \times \Omega_2)\}\end{aligned}$$

for every gamble f on $\Omega_1 \times \Omega_2$.

One way of building independent products is by means of the following condition:

Definition 4.119 *A coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega_1 \times \Omega_2)$ is called* factorising *when*

$$\underline{P}(fg) = \underline{P}(f\underline{P}(g)) \quad \forall f \in \mathcal{L}^+(\Omega_1), g \in \mathcal{L}(\Omega_2)$$

and

$$\underline{P}(fg) = \underline{P}(g\underline{P}(f)) \quad \forall f \in \mathcal{L}(\Omega_1), g \in \mathcal{L}^+(\Omega_2).$$

Both the independent natural extension and the strong product are factorising. Indeed, it can be proven [52, Theorem 28] that any factorising $\underline{P}$ is an independent product of its marginals, but the converse is not true. Under factorisation, it is not difficult to establish the following result.

Proposition 4.120 *Let $(\underline{F}_X, \overline{F}_X)$, $(\underline{F}_Y, \overline{F}_Y)$ be marginal p-boxes, and let $\underline{P}_X, \underline{P}_Y$ be their associated coherent lower previsions. Let $\underline{P}$ be a factorising coherent lower prevision on $\mathcal{L}(\Omega_1 \times \Omega_2)$ with these marginals. Then it induces the bivariate p-box $(\underline{F}, \overline{F})$ given by*

$$\underline{F}(x, y) = \underline{F}_X(x) \cdot \underline{F}_Y(y) \quad \text{and} \quad \overline{F}(x, y) = \overline{F}_X(x) \cdot \overline{F}_Y(y).$$

Proof It suffices to take into account that, if $\underline{P}$ is factorising, then

$$\underline{P}(A_{(x,y)}) = \underline{P}(I_{A_{(x,y^*)}} \cdot I_{A_{(x^*,y)}}) = \underline{P}(A_{(x,y^*)}) \cdot \underline{P}(A_{(x^*,y)}) = \underline{F}_X(x) \cdot \underline{F}_Y(y),$$

and similarly using conjugacy we deduce that

$$\overline{P}(A_{(x,y)}) = \overline{P}(A_{(x,y^*)} \cdot A_{(x^*,y)}) = \overline{P}(A_{(x,y^*)}) \cdot \overline{P}(A_{(x^*,y)}) = \overline{F}_X(x) \cdot \overline{F}_Y(y),$$

taking into account in the application of the factorisation condition that both gambles $A_{(x,y^*)}, A_{(x^*,y)}$ are positive, and recalling also that x^*, y^* denote the maxima of Ω_1, Ω_2, respectively. ∎

From this, it is easy to deduce that the p-box $(\underline{F}, \overline{F})$ induced by a factorising $\underline{P}$ is the lower envelope of the set of bivariate distribution functions

$$\{F : F(x, y) = F_X(x) \cdot F_Y(y) \text{ for } F_X \in (\underline{F}_X, \overline{F}_X), F_Y \in (\underline{F}_Y, \overline{F}_Y)\}.$$

In other words, the bivariate p-box can be obtained by applying the imprecise version of Sklar's theorem (Proposition 4.111) with the product copula.

In particular, this also holds for other (stronger) conditions than factorisation also discussed in [52], such as the Kuznetsov property.

Note also that in our definition of the marginal coherent lower previsions $\underline{P}_X, \underline{P}_Y$ we have considered the natural extensions of their restrictions to cumulative sets; however, the result still holds if we consider any other coherent extension, since in our use of the factorisation condition only the values in $A_{(x,y)}, A_{(x,y)}$ matter. We conclude then that, even if the independent natural extension and the strong product do not coincide in general [205, Section 9.3.4], they agree with respect to their associated bivariate p-box.

Interestingly, not all independent products induce the same p-box determined by the copula of the product:

Example 4.121 *Consider $\Omega_1 = \Omega_2 = \{0,1\}$ and let $\underline{F}_X = \underline{F}_Y$ be the marginal distribution functions given by $\underline{F}_X(0) = \underline{F}_Y(0) = 0.5$, $\underline{F}_X(1) = \underline{F}_Y(1) = 1$. They induce the marginal coherent lower previsions $\underline{P}_X, \underline{P}_Y$ given by*

$$\underline{P}_X(f) = \min\{f(0), 0.5f(0) + 0.5f(1)\} \text{ and } \underline{P}_Y(g) = \min\{g(0), 0.5g(0) + 0.5g(1)\}$$

for every $f \in \mathcal{L}(\Omega_1), g \in \mathcal{L}(\Omega_2)$. Their strong product is given by:

$$\underline{P}_X \boxtimes \underline{P}_Y := \min\{(0.25, 0.25, 0.25, 0.25), (0.5, 0, 0.5, 0), (0.5, 0.5, 0, 0), (1, 0, 0, 0)\}, \quad (4.20)$$

wherein the above equation a vector (a, b, c, d) is used to denote the vector of probabilities $\{(P(0,0), P(0,1), P(1,0), P(1,1))\}$. Let $\underline{P}$ be the coherent lower prevision given by

$$\underline{P} := \min\{(0.375, 0.125, 0.375, 0.125), (0.375, 0.375, 0.125, 0.125), (1, 0, 0, 0)\}.$$

Then the marginals of $\underline{P}$ are also $\underline{P}_X, \underline{P}_Y$. Moreover, we see from Equation (4.20) that $\underline{P}$ dominates $\underline{P}_X \boxtimes \underline{P}_Y$, and this allows us to deduce that $\underline{P}$ is weakly coherent with both $\underline{P}_X(\cdot|\Omega_2), \underline{P}_Y(\cdot|\Omega_1)$: given a gamble f on $\Omega_1 \times \Omega_2$,

$$\underline{P}(G(f|\Omega_2)) \geq (\underline{P}_X \boxtimes \underline{P}_Y)(G(f|\Omega_2)) \geq 0,$$

whence in particular $\underline{P}(G(f|y)) = \underline{P}(G(fI_y|\Omega_2)) \geq 0$ for every $y \in \Omega_2$. And since $\underline{P}_Y$ is the marginal of $\underline{P}$, it follows that we must have $\underline{P}(G(f|y)) = 0$: if it were $\underline{P}(G(f|y)) > 0$ then we would define the gamble g by $g(x,y) = f(x,y)$ and

$$0 = \underline{P}(g - \underline{P}_X(g)) \geq \sum_{y \in \Omega_2} \underline{P}(G(g|y)) > 0,$$

a contradiction. Similarly, $\underline{P}(G(f|\Omega_1)) \geq 0$ and $\underline{P}(G(f|x)) = 0$ for every $x \in \Omega_1$. Applying [137, Theorem 1], we conclude that $\underline{P}, \underline{P}_X(\cdot|\Omega_1), \underline{P}_Y(\cdot|\Omega_1)$ are weakly coherent, and since $\underline{P}_X(\cdot|\Omega_2), \underline{P}_Y(\cdot|\Omega_1)$ are coherent because they are jointly coherent with $\underline{P}_X \boxtimes \underline{P}_Y$, we deduce from the reduction theorem [205, Theorem 7.1.5] that $\underline{P}, \underline{P}_X(\cdot|\Omega_2), \underline{P}_Y(\cdot|\Omega_1)$ are coherent. Thus, $\underline{P}$ is an independent product. Its associated distribution function is given by

$$\underline{F}(0,0) = 0.375, \underline{F}(0,1) = 0.5, \underline{F}(1,0) = 0.5, \underline{F}(1,1) = 1.$$

This differs from the bivariate distribution function $\underline{F}$ induced by $\underline{P}_X \boxtimes \underline{P}_Y$, which is the product of its marginals and which satisfies therefore $\underline{F}(0,0) = 0.25$.

4.3.3 The roleof imprecise copulas in the imprecise orders

Next we study how imprecise copulas can be used to express the relationship between imprecise sto chastic dominance and statistical preference, that arise by using $_{FSD}$ and $_{SP}$ as the binary relation in Section 4.1. Afterwards, we shall stu dy the role of imprecise copulas with resp ect to imprecise bivariate sto chastic orders.

Univariate orders

We have seen in Section 3.2 that, although first degree sto chastic dominance do es not imply statistical preference ingeneral (seeExample3.43), thereare situations inwhich the imp lication holds (see Theorem 3.64), in termsof themarginal distributions of the variables and thecopula that determines their joint distribution.

Given two random variables X and Y, let usdenote by $C_{X,Y}$ the setof copulas that make sto chastic dominance imply statistical preferenceSince the latter dep ends on the joint distribution of the random variables, it may be that X is preferred to Y when theirjointdistributionis determined byacopula C_1 and Y ispreferred to X whenit is determined by different copula C_2.

In the imprecise framework, it is p ossible to establ ish the following conn ection b etween the imprecise sto chastic dominance and statistical preferenceWe shallassume that we have imprecise information ab out the marginal distributions (that we mo del by means of p-b oxes) and by the copula that links the marginal distributionsinto a joint (that we mo del by means of a set of copulas), in a manner similar to Prop osition 4.111:

Prop osition 4.122 *Considera coherent lowerprevision $\underline{P}$ definedon thespaceproduct $\mathcal{X} \times \mathcal{Y}$ of two finite spacesthat is factorising. Denote by $(\underline{F}, \overline{F})$ its associated bivariatep-box, that fromProposition 4.120 isbuilt fromthemarginal p-boxes$(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$ using the product copula. Then,it holds that:*

$$(\underline{F}_X, \overline{F}_X) \quad {}_{FSD_i} \quad (\underline{F}_Y, \overline{F}_Y) \qquad \mathcal{X} \quad {}_{SP_i} \quad \mathcal{Y}$$

for any $i = 1, \ldots, 6$,where $\mathcal{X}$ (respectively$\mathcal{Y}$) denotes the set of random variables whose cumulative distribution function belongs to $(\underline{F}_X, \overline{F}_X)$ ($(\underline{F}_Y, \overline{F}_Y)$, respectively).

Pro of We know from Prop osition 4.120 that $(\underline{F}, \overline{F})$ is built by applying the pro duct copula to their marginal p-b oxes.

- i=1 : Weknowthat forany F_X $(\underline{F}_X, \overline{F}_X)$ and F_Y $(\underline{F}_Y, \overline{F}_Y)$, F_X $_{FSD}$ F_Y. Since they are coupled by the pro duct copula, Theorem 3.44 implies P_{F_X} $_{SP}$ P_{F_Y}. Thus, $\mathcal{X}$ $_{SP_1}$ $\mathcal{Y}$.

- $i=2$: We know that there is F_X $(\underline{F}_X, \overline{F}_X)$ such that F_X $_{FSD}$ F_Y for any F_Y $(\underline{F}_Y, \overline{F}_Y)$. Since they are coupled by the pro duct copula, Theorem 3.44 implies P_{F_X} $_{SP}$ P_{F_Y} for any F_Y $(\underline{F}_Y, \overline{F}_Y)$. Then, X $_{SP_2}$ Y.
- $i=3$: We know that for any F_Y $(\underline{F}_Y, \overline{F}_Y)$ there is F_X $(\underline{F}_X, \overline{F}_X)$ such that F_X $_{FSD}$ F_Y. Then, for any P_{F_Y}, there is a P_{F_X} such that F_X $_{FSD}$ F_Y, and consequently, the pro duct copula links them, and by Theorem 3.44, P_{F_X} $_{SP}$ P_{F_Y}.
- $i=4$: We know that there are F_X $(\underline{F}_X, \overline{F}_X)$ and F_Y $(\underline{F}_Y, \overline{F}_Y)$ such that F_X $_{FSD}$ F_Y. Then, consider P_{F_X} and P_{F_Y}. Since they are coupled by the pro duct copula, Theorem 3.44 implies P_{F_X} $_{SP}$ P_{F_Y}.
- The pro of of cases $i=5$ and $i=6$ are similar to the one of cases $i=2$ and $i=3$. ∎

Remark 4.123 *Although we may think that the previous result also holds when we build the joint bivariate p-box from the marginal p-boxes by means of a set of copulas $\mathcal{C}$ $C_{X,Y}$, in the manner of Proposition 4.111, such a result does not seem to hold in general. The reason is that, as soon as one of the marginal p-boxes is imprecise (i.e., if its lower and the upper bounds do not coincide), we can find a distribution function inside the p-box associated with a neither continuous nor discrete random variable, and then, taking into account Theorem 3.64, we cannot assure the implication* $_{FSD}$ $_{SP}$ *unless we assume independence between the two p-boxes.*

Bivariate orders

As we saw in Equation (2.6), univariate sto chastic dominance can be expressed in terms of the comparison of exp ectations. It is also well-known that sto chastic dominance can b e expressed by means of the comparison of the survival distribution functions: given two random variables X and Y, their dis tribution functions are given by F_X and F_Y, and let $\overline{F}_X(t) = P(X > t)$ and $\overline{F}_Y = P(Y > t)$ denote their asso ciated survival distribution functions. Then, it holds that:

$$F_X(t) = P(X \leq t) \leq P(Y \leq t) = F_Y(t) \qquad \overline{F}_X(t) = 1 - F_X(t) \geq 1 - F_Y(t) = \overline{F}_Y. \tag{4.21}$$

Indeed, according to Equation (2.5), we have the following characterisations for first degree sto chastic dominance:

$$\begin{aligned} X \;_{FSD}\; Y \quad & F_X(t) \leq F_Y(t) \text{ for any } t \\ & E[u(X)] \geq E[u(Y)] \text{ for any increasing } u \\ & \overline{F}_X(t) \geq \overline{F}_Y(t) \text{ for any } t. \end{aligned}$$

In the bivariate cas e, the survival distribution functions are not related to the distribution functions as in Equation (4.21), since $P(X > t_1, Y > t_2) = 1 - P(X \leq t_1, Y \leq t_2)$. Then, these three conditions are not equivalent, and they generate three different sto chastic orders:

Definition 4.124 *Let (X_1, X_2) and (Y_1, Y_2) be two random vectors with bivariate distribution functions F_{X_1,X_2} and F_{Y_1,Y_2}. We say that:*

- (X_1, X_2) sto chastically dominates (Y_1, Y_2), *and denote it* $(X_1, X_2) \succeq_{FSD} (Y_1, Y_2)$, *if* $E[u(X_1, X_2)] \geq E[u(Y_1, Y_2)]$ *for any increasing* $u: \mathbb{R}^2 \to \mathbb{R}$.
- (X_1, X_2) *is preferred to* (Y_1, Y_2) with resp ect to the upp er orthant order, *and denote it* $(X_1, X_2) \succeq_{uo} (Y_1, Y_2)$, *if* $\overline{F}_{X_1,X_2}(t) \geq \overline{F}_{Y_1,Y_2}(t)$ *for any* $t \in \mathbb{R}^2$.
- (X_1, X_2) *is preferred to* (Y_1, Y_2) with res p ect to the lower orthant order, *and denote it* $(X_1, X_2) \succeq_{lo} (Y_1, Y_2)$, *if* $F_{X_1,X_2}(t) \leq F_{Y_1,Y_2}(t)$ *for any* $t \in \mathbb{R}^2$.

These three orders are equivalent in the univariate case, but not in th e bivariate. Next theorem describ e the relationships between these three orders:

Theorem 4.125 ([139, Theorem 3.3.2]) *If $X \succeq_{FSD} Y$, then $X \succeq_{lo} Y$ and $X \succeq_{uo} Y$. In addition, there is no implication between the lower and the upper orthant orders.*

In Remark 4.127 we will give an example where the lower and the upp er orthant orders are not equivalent.

Since any copula C is in particular a bivariate distribution function on $[0,1] \times [0,1]$ the previous orders can also b e ap plied to the comparison of copulas. Taking this into account, we can establis h the following result, that links the comparison of bivariate p-b oxes with the comparison of their asso ciated marginal p-b oxes.

Prop osition 4.126 *Let $(\underline{F}_{X_1}, \overline{F}_{X_1})$, $(\underline{F}_{X_2}, \overline{F}_{X_2})$, $(\underline{F}_{Y_1}, \overline{F}_{Y_1})$ be univariate p-boxes and $(\underline{F}_{Y_2}, \overline{F}_{Y_2})$ and the set of copulas $\mathcal{C}_X$ and $\mathcal{C}_Y$. Let $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$ be the bivariate p-boxes given by:*

$$(\underline{F}_X, \overline{F}_X) := \{C(F_{X_1}, F_{X_2}) : C \in \mathcal{C}_X, F_{X_1} \in (\underline{F}_{X_1}, \overline{F}_{X_1}), F_{X_2} \in (\underline{F}_{X_2}, \overline{F}_{X_2})\}$$
$$(\underline{F}_Y, \overline{F}_Y) := \{C(F_{Y_1}, F_{Y_2}) : C \in \mathcal{C}_Y, F_{Y_1} \in (\underline{F}_{Y_1}, \overline{F}_{Y_1}), F_{Y_2} \in (\underline{F}_{Y_2}, \overline{F}_{Y_2})\}.$$

Then, it holds that:

$$\left.\begin{array}{l} (\underline{F}_{X_1}, \overline{F}_{X_1}) \succeq_{FSD_i} (\underline{F}_{Y_1}, \overline{F}_{Y_1}) \\ (\underline{F}_{X_2}, \overline{F}_{X_2}) \succeq_{FSD_i} (\underline{F}_{Y_2}, \overline{F}_{Y_2}) \\ \mathcal{C}_X \succeq_{lo_i} \mathcal{C}_Y \end{array}\right\} \Rightarrow (\underline{F}_X, \overline{F}_X) \succeq_{lo_i} (\underline{F}_Y, \overline{F}_Y)$$

for $i = 1, \ldots, 6$.

Pro of

(i = 1) We knowthat:

$$F_{X_1} \quad (\underline{F}_{X_1}, \overline{F}_{X_1}), F_{Y_1} \quad (\underline{F}_{Y_1}, \overline{F}_{Y_1}), F_{X_1} \leq F_{Y_1}.$$
$$F_{X_2} \quad (\underline{F}_{X_2}, \overline{F}_{X_2}), F_{Y_2} \quad (\underline{F}_{Y_2}, \overline{F}_{Y_2}), F_{X_2} \leq F_{Y_2}.$$
$$C_X \quad C_X, C_Y \quad C_Y, C_X \leq C_Y.$$

Consider $F_X \quad (\underline{F}_X, \overline{F}_X)$ and $F_Y \quad (\underline{F}_Y, \overline{F}_Y)$. They can b e expres sed in the following way: $F_X(x, y) = C_X(F_{X_1}(x), F_{X_2}(y))$ and $F_Y(x, y) = C_Y(F_{Y_1}(x), F_{Y_2}(y))$. Then:

$$\begin{aligned} F_X(x, y) &= C_X(F_{X_1}(x), F_{X_2}(y)) \leq C_X(F_{Y_1}(x), F_{Y_2}(y)) \\ &\leq C_Y(F_{Y_1}(x), F_{Y_2}(y)) = F_Y(x, y). \end{aligned}$$

(i = 2) We know th at:

$$F_{X_1} \quad (\underline{F}_{X_1}, \overline{F}_{X_1}) \text{ such that } F_{X_1} \leq F_{Y_1} \quad F_{Y_1} \quad (\underline{F}_{Y_1}, \overline{F}_{Y_1}).$$
$$F_{X_2} \quad (\underline{F}_{X_2}, \overline{F}_{X_2}) \text{ such that } F_{X_2} \leq F_{Y_2} \quad F_{Y_2} \quad (\underline{F}_{Y_2}, \overline{F}_{Y_2}).$$
$$C_X \quad C_X \text{ such that } C_X \leq C_Y \quad C_Y \quad C_Y.$$

Consider $F_X(x, y) := C_X(F_{X_1}(x), F_{X_2}(y))$, and letusseethat $F_X \leq F_Y$ for any $F_Y(x, y) = C_Y(F_{Y_1}(x), F_{Y_2}(y))$:

$$\begin{aligned} F_X(x, y) &= C_X(F_{X_1}(x), F_{X_2}(y)) \leq C_X(F_{Y_1}(x), F_{X_2}(y)) \\ &\leq C_Y(F_{Y_1}(x), F_{X_2}(y)) = F_Y(x, y). \end{aligned}$$

(i =3) We know th at:

$$F_{Y_1} \quad (\underline{F}_{Y_1}, \overline{F}_{Y_1}), \quad F_{X_1} \quad (\underline{F}_{X_1}, \overline{F}_{X_1}) \text{ such that } F_{X_1} \leq F_{Y_1}.$$
$$F_{Y_2} \quad (\underline{F}_{Y_2}, \overline{F}_{Y_2}), \quad F_{X_2} \quad (\underline{F}_{X_2}, \overline{F}_{X_2}) \text{ such that } F_{X_2} \leq F_{Y_2}.$$
$$C_Y \quad C_Y \quad C_X \quad C_X \text{ such that } C_X \leq C_Y.$$

Consider $F_Y(x, y) = C_Y(F_{Y_1}(x), F_{Y_2}(y))$, andlet us check thatthere is F_X such that $F_X \leq F_Y$. We define $F_X(x, y) = C_X(F_{X_1}(x), F_{X_2}(y))$ such that $C_X \leq C_Y$, $F_{X_1} \leq F_{Y_1}$ and $F_{X_2} \leq F_{Y_2}$. Then:

$$\begin{aligned} F_X(x, y) &= C_X(F_{X_1}(x), F_{X_2}(y)) \leq C_X(F_{Y_1}(x), F_{Y_2}(y)) \\ &\leq C_Y(F_{Y_1}(x), F_{Y_2}(y)) = F_Y(x, y). \end{aligned}$$

(i =4) Weknow that:

$$F_{X_1} \quad (\underline{F}_{X_1}, \overline{F}_{X_1}), F_{Y_1} \quad (\underline{F}_{Y_1}, \overline{F}_{Y_1}) \text{ such that } F_{X_1} \leq F_{Y_1}.$$
$$F_{X_2} \quad (\underline{F}_{X_2}, \overline{F}_{X_2}), F_{Y_2} \quad (\underline{F}_{Y_2}, \overline{F}_{Y_2}) \text{ such that } F_{X_2} \leq F_{Y_2}.$$
$$C_X \quad C_X, C_Y \quad C_Y \text{ such that } C_X \leq C_Y.$$

Consider $F_X(x,y) = C_X(F_{X_1}(x), F_{X_2}(y))$ and $F_Y(x,y) = C_Y(F_{Y_1}(x), F_{Y_2}(y))$. It holds that $F_X \leq F_Y$:

$$F_X(x,y) = C_X(F_{X_1}(x), F_{X_2}(y)) \leq C_X(F_{Y_1}(x), F_{Y_2}(y))$$
$$\leq C_Y(F_{Y_1}(x), F_{Y_2}(y)) = F_Y(x,y).$$

($i = 5, i = 6$ **)** The pro of of thes e two cases is analogous to that of $i = 2$ and $i = 3$ resp ectively. ∎

Remark 4.127 *Note that under the hypotheses of Proposition 4.126 we do not necessarily have that* $(\underline{F}_X, \overline{F}_X) \preceq_{uo_i} (\underline{F}_Y, \overline{F}_Y)$. *To see this, consider the fol lowing probability mass functions (see [139, Example 3.3.3]):*

$X_2 \backslash X_1$	0	1	2
0	0	0	$\frac{1}{8}$
1	$\frac{1}{4}$	$\frac{1}{4}$	0
2	$\frac{1}{4}$	$\frac{1}{8}$	0

$Y_2 \backslash Y_1$	0	1	2
0	$\frac{1}{4}$	$\frac{1}{4}$	0
1	0	$\frac{1}{8}$	$\frac{1}{8}$
2	$\frac{1}{4}$	0	0

Then, $(X_1, X_2) \preceq_{lo} (Y_1, Y_2)$ *since* $F_{X_1,X_2} \leq F_{Y_1,Y_2}$. *However,* $(X_1, X_2) \not\preceq_{uo} (Y_1, Y_2)$, *since:*

$$\overline{F}_X(1,0) = P(X_1 > 1, X_2 > 0) = 0 < \frac{1}{8} = P(Y_1 > 1, Y_2 > 0) = \overline{F}_Y(1,0).$$

This example also shows that under the assumptions of Proposition 4.126 it does not necessarily hold that $(X_1, X_2) \preceq_{FSD} (Y_1, Y_2)$; *otherwise, we would deduce from Theorem 4.125 that* $(X_1, X_2) \preceq_{uo} (Y_1, Y_2)$, *a contradiction with the example above.*

A result similar to Prop osition 4.126 can be established when we consider the upp er instead of the lower orthant order:

Prop osition 4.128 *Let* $(\underline{F}_{X_1}, \overline{F}_{X_1})$, $(\underline{F}_{X_2}, \overline{F}_{X_2})$, $(\underline{F}_{Y_1}, \overline{F}_{Y_1})$ *be univariate p-boxes and* $(\underline{F}_{Y_2}, \overline{F}_{Y_2})$ *and the set of copulas* $\mathcal{C}_X$ *and* $\mathcal{C}_Y$. *Let* $(\underline{F}_X, \overline{F}_X)$ *and* $(\underline{F}_Y, \overline{F}_Y)$ *be the bivariate p-boxes given by:*

$$(\underline{F}_X, \overline{F}_X) := \{C(F_{X_1}, F_{X_2}) : C \in \mathcal{C}_X, F_{X_1} \in (\underline{F}_{X_1}, \overline{F}_{X_1}), F_{X_2} \in (\underline{F}_{X_2}, \overline{F}_{X_2})\}$$
$$(\underline{F}_Y, \overline{F}_Y) := \{C(F_{Y_1}, F_{Y_2}) : C \in \mathcal{C}_Y, F_{Y_1} \in (\underline{F}_{Y_1}, \overline{F}_{Y_1}), F_{Y_2} \in (\underline{F}_{Y_2}, \overline{F}_{Y_2})\}.$$

Then, it holds that:

$$\left.\begin{array}{r} (\underline{F}_{X_1}, \overline{F}_{X_1}) \preceq_{FSD_i} (\underline{F}_{Y_1}, \overline{F}_{Y_1}) \\ (\underline{F}_{X_2}, \overline{F}_{X_2}) \preceq_{FSD_i} (\underline{F}_{Y_2}, \overline{F}_{Y_2}) \\ \mathcal{C}_X \preceq_{uo_i} \mathcal{C}_Y \end{array}\right\} \Rightarrow (\underline{F}_X, \overline{F}_X) \preceq_{uo_i} (\underline{F}_Y, \overline{F}_Y),$$

for $i = 1, \ldots, 6$.

The pro of of this result is analogous to the one of Prop osition 4.126, and therefore omitted.

Natural extension and independent product

To conclude this section, we consider the particular cases where the bivariate p-boxes are made by means of the natural extension or a factorising product.

By Proposition 4.117, the natural extension of two marginal p-boxes $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$ is given by:

$$\underline{F}(x, y) = C_L(\underline{F}_X(x), \underline{F}_Y(y)) \text{ and } \overline{F}(x, y) = C_M(\overline{F}_X(x), \overline{F}_Y(y)). \tag{4.22}$$

This allows us to prove the following result:

Corollary 4.129 *Consider marginal p-boxes* $(\underline{F}_{X_1}, \overline{F}_{X_1})$, $(\underline{F}_{X_2}, \overline{F}_{X_2})$ *and* $(\underline{F}_{Y_1}, \overline{F}_{Y_1})$ *and* $(\underline{F}_{Y_2}, \overline{F}_{Y_2})$. *Let* $(\underline{F}_X, \overline{F}_X)$ *(respectively,* $(\underline{F}_Y, \overline{F}_Y)$*) denote the natural extension of the p-boxes* $(\underline{F}_{X_1}, \overline{F}_{X_1})$ *and* $(\underline{F}_{X_2}, \overline{F}_{X_2})$ *(respectively,* $(\underline{F}_{Y_1}, \overline{F}_{Y_1})$, $(\underline{F}_{Y_2}, \overline{F}_{Y_2})$*) by means of Equation* (4.22). *Then:*

$$\begin{array}{l} (\underline{F}_{X_1}, \overline{F}_{X_1}) \ \text{FSD}_i \ (\underline{F}_{Y_1}, \overline{F}_{Y_1}) \\ (\underline{F}_{X_2}, \overline{F}_{X_2}) \ \text{FSD}_i \ (\underline{F}_{Y_2}, \overline{F}_{Y_2}) \end{array} \qquad (\underline{F}_X, \overline{F}_X) \ \text{lo}_i \ (\underline{F}_Y, \overline{F}_Y)$$

for $i = 2, \ldots, 6$.

Proof The result follows immediately from Proposition 4.126. ∎

To see that the result does not hold for lo_i, consider the following example.

Example 4.130 *For* $j = 1, 2$, *let* $\underline{F}_{X_j} = \overline{F}_{X_j} = \underline{F}_{Y_j} = \overline{F}_{Y_j}$ *be the distribution function associated with a uniform distribution on* $[0, 1]$ *and let us denote it by* F. *Then, trivially:*

$$(\underline{F}_{X_j}, \overline{F}_{X_j}) \ \text{FSD}_1 \ (\underline{F}_{Y_j}, \overline{F}_{Y_j}) \text{ for } j = 1, 2.$$

To see that $(\underline{F}_X, \overline{F}_X) \ \text{lo}_1 \ (\underline{F}_Y, \overline{F}_Y)$, *it suffices to note that* $C_M(F, F)$ $(\underline{F}_X, \overline{F}_X)$ *and* $C_L(F, F)$ $(\underline{F}_Y, \overline{F}_Y)$, *and:*

$$C_M(F(0.5), F(0.5)) = C_M(0.5, 0.5) = 0.5 > 0 = C_L(0.5, 0.5) = C_L(F(0.5), F(0.5)).$$

We also saw in Proposition 4.120 that the bivariate p-box associated with a factorising coherent lower probability is obtained applying the product copula to the two marginal p-boxes. This fact allows us to simplify Propositions 4.126 and 4.128:

Corollary 4.131 *Consider two factorising coherent lower probabilities* $\underline{P}_X$ *and* $\underline{P}_Y$ *defined on* $X \times Y$, *where both sets are finite. Denote by* $(\underline{F}_X, \overline{F}_X)$ *and* $(\underline{F}_Y, \overline{F}_Y)$ *their associated bivariate p-boxes, that from Proposition 4.120 can be obtained by applying the product copula to their respective marginal distributions represented by the p-boxes*

$(\underline{F}_{X_1},\overline{F}_{X_1})$, $(\underline{F}_{X_2},\overline{F}_{X_2})$ *and* $(\underline{F}_{Y_1},\overline{F}_{Y_1})$ *and* $(\underline{F}_{Y_2},\overline{F}_{Y_2})$, *respectively. Then, it holds that:*

$$(\underline{F}_{X_1},\overline{F}_{X_1}) \ \text{FSD}_i\ (\underline{F}_{Y_1},\overline{F}_{Y_1}) \qquad (\underline{F}_{X},\overline{F}_{X}) \ \text{lo}_i\ (\underline{F}_{Y},\overline{F}_{Y})$$
$$(\underline{F}_{X_2},\overline{F}_{X_2}) \ \text{FSD}_i\ (\underline{F}_{Y_2},\overline{F}_{Y_2}) \qquad (\underline{F}_{X},\overline{F}_{X}) \ \text{uo}_i\ (\underline{F}_{Y},\overline{F}_{Y}).$$

Proof We have seen in Proposition 4.120 that the bivariate p-box associated with a factorising coherent lower probability is made by considering the product copula applied to the marginal p-boxes. Then, this result is a particular case of Propositions 4.126 and 4.128. ■

4.4 Applications

To conclude the chapter, we give some possible applications of the extension of stochastic orders to an imprecise framework. We start with two possible applications of imprecise stochastic dominance: the comparison of Lorenz Curves and that of cancer survival rates. Lorenz Curves are a well-known economic tool that measure how the wealth of a population is distributed. Since Lorenz Curves can be seen as distribution functions, we can compare them by means of stochastic dominance. Furthermore, in some cases the economical analysis is made for geographical regions that comprise several countries, like for example Nordic countries, Southern Europe, American, ... Then, we can use the imprecise stochastic dominance to compare the sets of Lorenz Curves associated with these groups of countries. On the other hand, some kind of cancer sites can also by grouped into Digestive, Respiratory, Reproductive or Other. Then, it is possible to compare the survival rates of the group of cancer by comparing their associated set of mortality rates, that can be expressed as distribution functions. Then, also the imprecise stochastic dominance could be applied.

Afterwards, we focus on a Multi-Criteria Decision Making problem, where it is possible to find imprecision in the utilities or in the beliefs. This allows us to illustrate how both the imprecise stochastic dominance and statistical preference can be used, as well as the strong and weak dominance introduced in Section 4.2.2.

4.4.1 Comparison of Lorenz curves

As we mentioned in Section 2.1.1, the notion of stochastic dominance has been applied in many different contexts. One of the most interesting is in the field of social welfare [3, 117, 190], for comparing *Lorenz curves*. They are a graphical representation of the cumulative distribution function of the wealth: the elements of the population are ordered according to it, and the curve shows, for the bottom x% elements, what percentage y% of

Country-year	0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1
Australia-1994	5.9	12.01	17.2	23.57	41.32
Canada-2000	7.2	12.73	17.18	22.95	39.94
China-2005	5.73	9.8	14.66	22	47.81
Finland-2000	9.62	14.07	17.47	22.14	36.7
FYR Macedonia-2000	9.02	13.45	17.49	22.61	37.43
Greece-2000	6.74	11.89	16.84	23.04	41.49
India-2005	8.08	11.27	14.94	20.37	45.34
Japan-1993	10.58	14.21	17.58	21.98	35.65
Maldives-2004	6.51	10.88	15.71	22.66	44.24
Norway-2000	9.59	13.96	17.24	21.98	37.23
Sweden-2000	9.12	13.98	17.57	22.7	36.63
USA-2000	5.44	10.68	15.66	22.4	45.82

Table 4.2: Quintiles of the Lorenz Curves ass o ciated with different countries.

the total wealth they have. Hence, the Lorenz curve can be used as a measure of equality: the closest the curve is to the straight line, the more equal the asso ciated society is.

Ifwe havethe Lorenzcurves oftwo differentcountries, we can compare them by means of sto chastic dominance: if one of them is domi nated by the other, the closest to the straight line will be asso ciated with a more equal so ciety,and will therefore be considered preferable. Inthis section, we are going to use our extensions of sto chastic dominance to compare sets of Lorenz curves asso ciated with countries in different areas of the world. We shall consider the Lorenz curves asso ciated with the qui ntiles of the empirical distribution functions. Table4.2 providesthe wealthineachofthequintiles (Source data: World Bankdatabase. http://timetric.c om/datas et/worldbank):

To make the comparison by means of the extensions of sto chastic dominance clearer, we are going to consider the cumulative distribution from the richest to the po orest group: in this way, we will always obtain a curve which is ab ove the straight line, and it will comply with our idea of considering preferable the smalle st distribution function. If we applythis to thedata in Table 4.2,weobtain the dataofTable4.3.

We are going to group thesecountriesby continents/regions:

- Group 1: China, Japan,India.
- Group 2: Finland, Norway,Sweden.
- Group 3: Canada, USA.
- Group 4: FYR Macedonia, Greece.

Country-year	F(0.2)	F(0.4)	F(0.6)	F(0.8)	F(1)
Australia-1994	41.32	64.89	82.09	94.1	100
Canada-2000	39.94	62.89	80.07	92.8	100
China-2005	47.81	69.81	84.47	94.27	100
Finland-2000	36.7	58.84	76.31	90.38	100
FYR Macedonia-2000	37.43	60.04	77.53	90.98	100
Greece-2000	41. 49	64.53	81.37	93.26	100
India-2005	45.34	65.71	80.65	91.92	100
Japan-1993	35.65	57.63	75.21	89.42	100
Maldives-2004	44.24	66.9	82.61	93.49	100
Norway-2000	37.23	59.21	76.45	90.41	100
Sweden-2000	36.63	59.33	76.9	90.88	100
USA-2000	45.82	68.22	83.88	94.56	100

Table 4.3: Cumulative distribution functions asso ciated with the Lorenz Curves of the countries.

	Group1	Group2	Group3	Group4	Group5
Group1	$\equiv FSD_{2,5}$	FSD_2	FSD_2	FSD_2	FSD_2
Group2	FSD_5	$\equiv FSD_{3,6}$	FSD_1	FSD_1	FSD_1
Group3	$\equiv FSD_4$		$\equiv FSD_{2,5}$	FSD_2	FSD_2
Group4	FSD_5		FSD_5	$\equiv FSD_{3,6}$	$FSD_{3,6}$
Group5	FSD_5		FSD_5		$\equiv FSD_{3,6}$

Table 4.4: Result of the comparison of the regions by means of the imprecise sto chastic dominance.

- Group 5: Australia, Maldives.

The relationships b etween thes e groups are summarised in Table 4.4.

This means for instance that the set of distribution functions in the first group strictly dominates the second group according to definition (FSD_2), while thesecond group strictly dominates the first group according to definition (FSD_5). This is because the b est country in the first group (Japan) sto chastically dominates all the countries in the second group, but the worst (China) is sto chastically dominated by all countries in the second group. This, together with Prop osition 4.3, implies that thefirst group strictly dominates the second accordingto (FSD_3), is strictly dominated by the second according to (FSD_6), thatthey areindifferentaccording to (FSD_4) and incomparable according to (FSD_1).

Similar considerations hold for the other pairwise comparisons.Forinstance, group

4 strictly dominates group 5 according to (FSD_3), (FSD_6), but it does not dominate it according to (FSD_2), (FSD_5). This also shows that conditions (FSD_2) and (FSD_3) are not equivalent (and similarly for (FSD_5) and (FSD_6)).

The cells where we have left a blank space mean that no dominance relationship is satisfied: for instance, group 3 does not dominate group 2 according to any of the definitions.

Since all the groups have more than one element, they will not satisfy (FSD_1) when comparing them to themselves. It follows from Remark 4.31 that they are always indifferent to themselves according to (FSD_3), (FSD_4) and (FSD_6); they are indifferent to themselves according to (FSD_2) when they have a best-case-scenario (as it is the case for groups 1 and 3), and indifferent according to (FSD_5) when they have a worst-case scenario (as it is the case again for groups 1 and 3), and incomparable according to these definitions in the other cases.

Note that we can also use the above data to illustrate some of the results in this paper: for instance, we saw in Remark 4.9 that condition (FSD_2) is transitive, and in the table above we see that group 1 is preferred to group 3 according to (FSD_2) and group 3 is preferred to group 4 according to (FSD_2): this allows us to infer immediately that group 1 is preferred to group 4 according to this condition. The comparison of the first two groups is an instance of Proposition 4.32, because the p-box induced by the first group is strictly more imprecise (i.e., it has a smaller lower cumulative distribution and a greater upper cumulative distribution function) than that of the second group.

Remark 4.132 *In economy, the Gini Index is a well-known inequality measure that express how the incomes of a population are shared. It takes values between 0 and 1, where a Gini Index of 0 means perfect equality for the incomes of the people, while a Gini Indez of 1 express a total inequality in the incomes. Thus, the greater the Gini Index is, the more inequality the incomes of a population are.*

The Gini Index is quite related to Lorenz curves: given a Lorenz Curve F, that express the distribution function of the wealth of a population (a country, a region, ...), its associated Gini index is defined by:

$$G = 2\int_0^{100} (x - F(x))dx.$$

Thus, the closer the Lorenz curve is to the straight $y=x$, the smaller the Gini index is.

In the imprecise framework, if we are working with a p-box that represents the Lorenz curve, we can compute the lower and the upper Gini Indexes, that are a lower and an upper bound of the Gini Index, simply by considering the Gini indexes of the upper and the lower bounds of the p-box. Then, for any p-box $(\underline{F}, \overline{F})$ representing Lorenz curve F we obtain a Gini index given in an interval form: $[\underline{G}, \overline{G}]$, where $\underline{G}$ is the Gini index associated with $\overline{F}$ and $\overline{G}$ is the Gini index associated with $\underline{F}$. Then, in order to compare *the Gini intervals associated with two imprecise Lorenz curves, it is possible to consider the usual orderings for real intervals (see for instance [69, 78]).*

4.4.2 Comparison ofcancer survival rates

According to [28], long-term cancer survival rates have substantially improved in the past decades.However, there are still some kinds of cancer whose survi val rates could clearly b e improved.Here, we use the survi val rates of different cancer sites given in [28].These can be group ed in Digestive, Respiratory, Repro ductive and Other, and we shall compare the survival rates of these typ es applying imprecise sto chastic dominance.

Table 4.5showsthesurvivalrates of different cancer sites (see [28]).

Note that it is possible to transform the survival rates of Table 4.5 into cumulative distribution functions. In thiscase, we assum e the distribution functions to b e defined in the interval [0,100], and we imp ose the condition $F(100) = 1$, that means that the survivalrate after 100years ofthe cancer diagnostic is zero. The resu lts are showed in Table 4.6.

These cancer sites can be group ed as follows:

Digestives Colon (C), Rec tum (R), Oral cavity and pharynx(OCP), Stomach (S), Oesophagus(O), Liver and intrahepaticbileduct (LIBD), Pancreas (P).

Respiratory Larynx (L), Lung and bronchus (LB).

Repro ductive Prostate (Pr), Testis (T), Breast (B), Cervix uteri (CU), Corpus uteri and uterus (CUU), Ovary (Ov).

Other Melanomas (M), Urinary bladder (UB), Kidney and renal pelvis (KRP), Brain and other nervous system (BNS), Thyroid (Th), Ho dgkin's disease (HD), Non-Ho dgkin lymphomas (NHL), Leukaemias (L).

Let us compare these kinds of cancer by means of the imprecise sto chastic dominance. Note that in this case, give n two distribution functions F_1 and F_2 thatrepresent the mortality rates of two cancer sites, F_1 $_{FSD}$ F_2 meansthatthe cancer F_1 isless deadly than th e cancer F_2, or equivalently, that the cancer F_1 has a greater survivalrate than the cancer F_2.

First of all, note that Pancreas (P) is the worst cancer with resp ect to sto chastic dominance, since $F < F_P$ for anyother distribution function F. This impliesthat Digestive is $_{FSD}$ 5 dominated by the other three groups, and then, from a p essimistic point of view, digestive cancersarethe worst. Furthermore, ProstateandThyroidcancersareless deadly than any of the digestive cancers, and then both Repro ductive and Other groups

	Relative survivalrate,%			
	1year	4years	7years	10 years
Cancer site				
Colon	80.7	65.6	60.5	58.2
Rectum	86.3	68.2	61.2	57.9
Oral cavity and pharynx	82.9	63.0	56.1	50.2
Stomach	49.0	27.0	22.9	20.8
Oesophagus	43.4	17.9	13.8	11.8
Liver and intrahepaticbile duct	34.5	15.2	11.0	9.2
Pancreas	23.0	6.2	4.5	3.8
Larynx	85.9	66.3	57.0	49.6
Lung and bronchus	41.2	17.5	13.0	10.5
Prostate*	99.6	98.6	97.9	97.0
Testis*	97.8	95.7	95.4	95.0
Breast**	97.5	90.4	85.8	82.6
Cervix uteri**	88.0	72.3	68.3	66.1
Corpus uteri anduterus**	92.4	83.9	81.5	80.3
Ovary**	74.9	48.5	38.8	35.0
Melanomas	97.3	92.2	90.3	89.5
Urinary bladder	90.1	80.9	76.4	72.7
Kidney and renal pelvis	80.8	69.3	63.8	59.4
Brain and othe r nervous system	56.4	35.1	30.6	27.9
Thyroid	97.6	96.9	96.3	95.9
Ho dgkin's disease	92.4	85.8	82.2	79.6
Non-Ho dgkin lymphomas	77.2	65.1	59.0	54.3
Leukae mias	70.2	55.0	48.3	43.8

Table 4.5: Estimationof relativesurvivalratesby cancersite. The ratesarederivedfrom SEER 1973-98 database,all ethnic groups, both sexes (except (*), only formale, and (**) for female). [191].

FSD_2 dominates Digestive. However,Digestiveand Respiratoryare incomparablewith resp ect to(FSD_2) and (FSD_3), and they are equivalent with resp ect to (FSD_4),since $F_P > F_{LB} > F_C$. Also Digestive is(FSD_4) equivalent to Repro ductive and Other groups, since $F_P > F_{Ov} > F_C$ and $F_P > F_{BNS} > F_C$.

Since Lung and Brounch cancer has a greater mortality than any Repro ductive cancer, Respiratory is FSD_5 dominated by Repro ductive group. Furthermore , they are not comparable with resp ect to (FSD_2) and indifferent with resp ect to (FSD_4) since $F_L < F_{Ov} < F_{LL}$.

Finally, since Brain and other nervous system cancer is sto chastically dominated by any Reproductive cancer, Repro ductive FSD_5 dominatesOther group, andthey are

	Cumulative distribution functions			
	$F(1)$	$F(4)$	$F(7)$	$F(10)$
Cancer site				
Colon	0.193	0.344	0.395	0.418
Rectum	0.137	0.318	0.388	0.421
Oral cavity and pharynx	0.171	0.370	0.439	0.498
Stomach	0.510	0.730	0.771	0.792
Oesophagus	0.566	0.821	0.862	0.882
Liver an d intrahepatic bile duct	0.655	0.846	0.890	0.908
Pancreas	0.770	0.938	0.955	0.962
Larynx	0.141	0.337	0.430	0.504
Lung and bronchus	0.588	0.825	0.870	0.895
Prostate	0.004	0.014	0.021	0.030
Testis	0.022	0.043	0.046	0.050
Breast	0.025	0.096	0.142	0.174
Cervix uteri	0.120	0.277	0.317	0.339
Corpus uteri anduterus	0.076	0.161	0.185	0.197
Ovary	0.251	0.515	0.612	0.650
Melanomas	0.027	0.078	0.097	0.105
Urinary bladder	0.099	0.191	0.236	0.273
Kidney and renal pelvis	0.192	0.307	0.362	0.406
Brain and othe r nervous system	0.436	0.649	0.694	0.721
Thyroid	0.024	0.031	0.037	0.041
Ho dgkin's disease	0.076	0.142	0.178	0.204
Non-Ho dgkin lymphomas	0.228	0.349	0.410	0.457
Leukaemias	0.298	0.450	0.517	0.562

Table 4.6: Estimation ofrelativemortalityratesby cancersite.

equivalent with resp ect to (FSD_4) since $F_M < F_{CU} < F_{BNS}$.

The results aredepicted in Table 4.7.

Thus, according to our results, Digestive cancer seems to be the group with a greater mortality rate, while Repro ductive cancer seems to b e the least deadly.

4.4.3 Multiattributedecision making

In this section, we shall illustrate the extension of statistical preference to acontext of im precision by means of an application to decision making. We shall consider two different scenarios: on the one hand, we sh allcompare two alternatives in acontext of

	Digestive	Respiratory	Repro ductive	Other
Digestive	$\equiv$ FSD 5	$\equiv$ FSD 4	$\equiv$ FSD 4	$\equiv$ FSD 4
Respiratory	FSD 5	$\equiv$ FSD 2,5		$\equiv$ FSD 4
Repro ductive	FSD 2,5	FSD 5	$\equiv$ FSD 5	FSD 5
Other	FSD 2,5	FSD 2	$\equiv$ FSD 4	$\equiv$ FSD 2,5

Table 4.7: Result of the comparis on of the different groups of cancer by means of the imprecise sto chastic dominance.

imprecise information ab out their utilities or probabilities, by means of the results in Sections 4.2.1 and 4.2.2; on the other hand, we shall consider the comparison of two sets of alternatives, by means of the techniques established in Section 4.1. Our runn ing example throughout this section is based on [118, Section 4].

A decision problem with uncertain beliefs

Consider a decision problem where we must cho ose b etween n alternatives $a_1, \dots, a_n$, whose rewards dep end on the values of the states of nature $\theta_1, \dots, \theta_m$, which hold with certain probabilities $P(\theta_1), \dots, P(\theta_m)$.

Let us start by assuming that there is uncertainty ab out these probabilities, that we mo del by means of a set of probability measures $\mathcal{P}$. Then, we shall compare any two alternatives by means of the concepts of weak and strong $\mathcal{P}$-preference we have considered in Section 4.2.2.

Example 4.133 *A company must choose where to invest its money. The alternatives are: a_1-a computer company; a_2-a car company; a_3-a fast food company. The rewards associated with the investment depend on an attribute c_1: "economic evolution", which may take the values θ_1-"very good", θ_2-"good", θ_3-"normal" or θ_4-"bad". The probabilities of each of these states are expressed by means of an interval. The rewards associated with any combination (alternative, state) are expressed in a linguistic scale, with values $S = \{s_0, s_1, s_2, s_3, s_4, s_5, s_6\}$ (very poor, poor, slightly poor, normal, slightly good, good, very good). The available information is summarised in the fol lowing table:*

	θ_1	θ_2	θ_3	θ_4
	[0 .1, 0 .4]	[0.2 , 0.7]	[0.3, 0.4]	[0.1 , 0.4]
a_1	s_4	s_3	s_3	s_2
a_2	s_5	s_4	s_4	s_2
a_3	s_2	s_3	s_5	s_4

Hence, the set $\mathcal{P}$ of probability measures for our beliefs is given by

$$\mathcal{P} = \{(p_1, p_2, p_3, p_4) : p_1 + p_2 + p_3 + p_4 = 1,$$
$$p_1 \in [0.1, 0.4], p_2 \in [0.2, 0.7], p_3 \in [0.3, 0.4], p_4 \in [0.1, 0.4]\}.$$

Since the rewards are expressed in a qualitative scale, we are going to compare the different alternatives by means of statistical preference. We obtain that:

$$Q(a_1, a_2) = \tfrac{1}{2} p_4 \in [0.05, 0.2].$$
$$Q(a_1, a_3) = p_1 + \tfrac{1}{2} p_2 \in [0.2, 0.5].$$
$$Q(a_2, a_3) = p_1 + p_2 \in [0.3, 0.6].$$

We deduce that, using statistical preference as our basic binary relation:

- $a_2 \succ^{\mathcal{P}}_{s} a_1$ *and* $a_2 \succ^{\mathcal{P}}_{w} a_1$.
- $a_3 \succ^{\mathcal{P}}_{s} a_1$ *and* $a_3 \equiv^{\mathcal{P}}_{w} a_1$.
- $a_2 \equiv^{\mathcal{P}}_{w} a_3$ *and they are incomparable with respect to strong $\mathcal{P}$-preference.*

Consequently, with respect to the strong preference, the car company is preferred to the computer company, while the car and the fast food company are incomparable. With respect to weak preference, the car company is also preferred to the computer company, while the fast food company is indifferent to the car and the computer companies.

A decision problem with uncertain rewards

Assume next that we have precise information about the probabilities of the different states of nature but that we have imprecise information about the utilities associated with the different rewards. Let us model this case by means of a random set, as we discussed in Section 4.2.1.

Example 4.133 (Cont) *Assume that the probability of the different states of nature is given by:*

$$P(\theta_1) = 0.2 \quad P(\theta_2) = 0.25 \quad P(\theta_3) = 0.3 \quad P(\theta_4) = 0.25,$$

but that we cannot determine precisely the consequences associated with each combination (alternative, state). We model the available information by means of a set of possible consequences, that we summarise in the following table:

	θ_1	θ_2	θ_3	θ_4
	0.2	0.25	0.3	0.25
a_1	$[s_4, s_5]$	$\{s_3\}$	$[s_2, s_3]$	$\{s_2\}$
a_2	$\{s_5\}$	$[s_3, s_4]$	$[s_3, s_5]$	$[s_2, s_4]$
a_3	$\{s_2\}$	$[s_3]$	$[s_3, s_5]$	$[s_3, s_4]$

Since again we have qualitat ive rewards,we shall use statistical preference to compare the different alternatives. Taking into account that the utility space is finite, we deduce from Proposition 4.78 that the comparison of the random set s associated with each of the alternatives reduces to the comparison of their maximaand minima measurable selections. Moreover,since the utilityspace is finite, $_{SP_2}$ $_{SP_3}$ *and* $_{SP_5}$ $_{SP_6}$.

Let us compare alternatives a_1, a_2*:*

$$Q(\min a_1, \max a_2) = 0.$$
$$Q(\min a_1, \min a_2) = 0.25.$$
$$Q(\max a_1, \max a_2) = 0.1.$$
$$Q(\max a_1, \min a_2) = 0.5.$$

Using Proposition 4.78, we conclude that a_2 $_{SP_i}$ a_1 *for* $i = 1, 2, 3, 5{,}6$ *and* $a_1 \equiv_{SP_4} a_2$.

With respect to alternatives a_1 *and* a_3*, we obtain that:*

$$Q(\min a_1, \max a_3) = 0.325$$
$$Q(\min a_1, \min a_3) = 0.325.$$
$$Q(\max a_1, \max a_3) = 0.325.$$
$$Q(\max a_1, \min a_3) = 0.475.$$

UsingProposition 4.78, we concludethat a_3 $_{SP_i}$ a_1 *for* $i{=}4$ *andasa consequence also for* $i = 1{,}2, 3, 5{,}6$.

Final ly, if we compare alternatives a_2 *and* a_3*, we obtain that:*

$$Q(\min a_2, \max a_3) = 0.325$$
$$Q(\min a_2, \min a_3) = 0.475.$$
$$Q(\max a_2, \max a_3) = 0.725.$$
$$Q(\max a_2, \min a_3) = 1.$$

UsingProposition 4.78, weconclude that a_2 $_{SP_i}$ a_3 *for* $i = 2{,}3$, a_3 $_{SP_i}$ a_2 *for* $i = 5{,}6$, $a_2 \equiv_{SP_4} a_3$ *and they areincomparable with respect to* $_{SP_1}$. *Hence, inthiscasethechoice between* a_2 *and* a_3 *woulddependonourattitudetowardsrisk, which woulddetermineif wefocus onthe best ortheworst-casescenarios. Consequently, boththecarandthefast foodcompaniesarepreferred tothecomputer one. However, thepreferencebetweenthe car and fast foodcompanies dependson the chosencriteria.*

A decision problem between sets of alternatives

Assume now that we have precise b eliefs and utilities but the choice must be made between sets of alternatives instead of pairs. Inthatcase, weshallapplytheconditions andresults from Section 4.1.

Example 4.133 (Cont) *Assume now that we may invest our money in another company a_4 in the telecommunications area, and that the choice must be made between two portfolios: one –that we shall denote X–made by alternatives a_1, a_2, and another –denoted by Y–made by a_3, a_4. Assume that the rewards associated with each alternative are given by the following table:*

	θ_1	θ_2	θ_3	θ_4
	0.2	0.25	0.3	0.25
a_1	75	60	55	50
a_2	80	65	55	40
a_3	60	55	50	55
a_4	80	55	40	65

where the utilities are now expressed in a $[0, 100]$ scale.

If we compare these alternatives by means of stochastic dominance, we obtain that $a_1 \ \mathrm{FSD}\ a_3$, $a_2 \ \mathrm{FSD}\ a_4$ and any other pair (a_i, a_j) with $i \{1,2\}, j \{3,4\}$ are incomparable with respect to stochastic dominance. Hence, $X \ \mathrm{FSD}_i\ Y$ for $i = 3, 4, 6$ and they are incomparable with respect to FSD_i for $i = 1, 2, 5$.

Note that this example is an instance where FSD_2 is not equivalent to FSD_3 and FSD_5 is not equivalent to FSD_6, because there is neither a maximum nor a minimum in the sets of distribution functions associated with X, Y.

On the other hand, if we compare any two alternatives by means of statistical preference, we obtain the following profile of preferences:

$$Q^{X,Y} := \begin{matrix} 0.75 & 0.55 \\ 0.75 & 0.65 \end{matrix}.$$

Using Remark 4.73, we obtain that $X \ \mathrm{SP}_1\ Y$, and as a consequence $X \ \mathrm{SP}_i\ Y$ for $i = 2, \ldots, 6$ and also $X \ \mathrm{SP}_{mean}\ Y$. Hence, from the point of view of statistical preference the first portfolio should be preferred to the second.

4.5 Conclusions

In this chapter we have considered the comparison of alternatives under both uncertainty and imprecision. As in Chapter 3, alternatives defined under uncertainty have been modelled by means of random variables, while the imprecision about the random variables has been modelled with sets of random variables, or in a more general situation, imprecise probability models.

We have extended binary relations to the comparison of sets of random variables instead of pairs of them. For this aim, we considered six possible generalisations. We have seen that the interpretation of each extension is related to the extensions of expected utility within imprecise probabilities.

We have mainly focused on two stochastic orders in this report: stochastic dominance and statistical preference. When we consider the binary relation to be first degree stochastic dominance, its extensions are related to the comparison of the p-boxes associated with the sets of random variables to compare. Also, according to the usual characterisation of stochastic dominance in terms of the comparison of the expectation of the increasing transformations of the random variables, we can also relate imprecise stochastic dominance to the comparison of the upper or lower expectations of the increasing transformation of the sets of random variables. We have also seen that our approach to extend stochastic dominance to the comparison of sets of random variables includes Denoeux approach ([61]) as a particular case, and we have also applied stochastic dominance to the comparison of possibility measures.

The extension of statistical preference has been connected to the comparison of the lower and upper medians of some set of random variables. We have seen that, when the sets of random variables to compare are finite, their comparison can be made by means of the pointwise comparison of the random variables by means of statistical preference, aggregating them with an aggregation function, and we have showed that the six extensions of statistical preference can be expressed in terms of aggregation functions.

We have also investigated two situations which can be considered as particular cases of the comparison of sets of random variables. On the one hand, we considered two random variables with imprecision on the utilities. That is, imprecise knowledge about the value of $X(\omega)$ and $Y(\omega)$. To model this imprecision, we have considered random sets Γ_X and Γ_Y, with the interpretation that the real value of $X(\omega)$ (respectively, $Y(\omega)$) belongs to $\Gamma_X(\omega)$ (respectively $\Gamma_Y(\omega)$). Then, we know that the random variables X, Y to be compared belong to the set of measurable selections of the random sets. Thus, the comparison of the random variables with imprecise utilities is made by the comparison of the random sets, which in fact can be made by means of the comparison of their associated sets of measurable selections.

On the other hand, we have also considered two random variables defined in a probability space whose probability is imprecisely described. We modelled this lack of information by means of a credal set. Then the random variables depend on the exact probability of the initial space. To deal with this imprecision we have introduced two new definitions: strong and weak preference.

We have seen that some binary relations, such as statistical preference, depend on the joint distribution of the random variables. In this framework Sklar's Theorem is a powerful tool that allows to build the joint distribution function from the marginals. However, there could be imprecision either in the marginal distributions, for example by considering p-boxes instead of distribution functions, or in the copula that links these marginals. For this reason we have developed a mathematical model that allows us to deal with this problem. In the first step, we showed that the infimum and supremum of sets of bivariate distribution functions are not bivariate distribution functions in general, because it may not satisfy the rectangle inequality. We have studied this problem by means of imprecise probabilities, extending the notion of p-box to the bivariate case. Then, the infimum and supremum of bivariate distribution functions determine a coherent lower probability that satisfies some imprecise version of the rectangle inequalities.

On the other hand we have considered the case where the lack of information lies in the copula that links the marginals. For this problem, we have extended copulas to the imprecise framework, and we have proven an imprecise version of the Sklar's Theorem. Finally, we have seen how bivariate p-boxes and this imprecise version of the Sklar's Theorem could be applied to one and two-dimensional stochastic orders.

Since in the real life it is common to encounter situations in which the information is imprecisely described, the results of this chapter have several applications. We have showed how imprecise stochastic dominance can be applied in the comparison of Lorenz Curves and cancer survival rates, and illustrated the usefulness of imprecise statistical preference for multicriteria decision making problems under uncertainty.

5 Comparison of alternatives underimprecision

Chapter 3 wasdevoted tothe comparison of alternatives in a decision proble m under acontext of uncertainty, where these alternatives were mo delled by means ofrandom variables. In C hapter 4 we added imprecisi on to the original problem, and we studied the comparison of sets of random variables. In thischapterwe shall assume that the alternatives are defined under imprecision but withoutuncertainty. In thiscasewe need not use probability theory, as the outcomes of the differe nt alternative will b e constant. However, the imprecision makes crisp sets not to an adequate mo del of the available information. Because of this, we shall use amore flexibletheorythanthe oneofcrisp sets: thatoffuzzysetsoranyofitsextensions, suchasthetheoryofIF-setsorIVF-sets.

While for the comparison of random variables or sets of rand om variables we use sto chastic orders, and some to ols of the imprecise probability theory, for the comparison of IF-sets orIVF-sets we shall use some measures of comparison of these kinds of sets.

In the framework of fuzzy set theory, we can find in the literature se veral measures of comparison between fuzzy sets. The more us ual measures of comparison are dissimilarities ([119]), dissimilitudes ([44]) and di vergences ([159]),inaddition to classical distances. Otherauthors, likeBouchon-Meunier([27])triedtodefinea generalmeasure of comparison between fuzzy sets,that includethe citedmeasures asparticular cases. The last attempt was made by C ou so et al.([45]) where someusual axioms requiredby the measures of comparison of fuz zy sets are collected and analyzed.

Montes ([159])made a completestudy of divergencesas comparisonmeasures of fuzzy sets.She intro duced a particular kind of divergences, the so-called lo cal divergences, which have been proved to be very useful.

Howe ver,in theframeworkofIF-sets, inthe literaturewecanonly finddistances for IF-se ts and a lot of examples of IF-dissimil arities (see for example [36, 37, 85, 89, 92, 111, 113, 114, 138, 193, 212]), butthere isnota thoroughmathematical theory of comparison of IF-sets.

For this reason, the first part of this chapter is devoted to the generalization of the comparison measures from fuzzy sets to IF-sets. Note that even though in this part we shall deal with IF-sets, our comments in Section 2.3 guarantee that all our results remain valid for IVF-sets.

Afterwards, we shall investigate the relationship between IF-sets and imprecise probabilities. In this second part, we shall interpret IF-sets as IVF-sets, because this allows for a clearer connection to imprecise probability. Thus, we shall assume that the IVF-set is defined on a probability space, and that it may be thus interpreted as a random set. Then, we shall investigate its main properties.

The results we present in this chapter have several applications. On the one hand, the measures of comparison of IF-sets have been used in several fields, such as pattern recognition ([92, 93, 94, 113, 114]) or decision making ([194, 211]), among others. On the other hand, the connection between IVF-sets and imprecise probabilities will be very useful when producing a graded version of stochastic dominance, and they shall allow us to propose a generalization of stochastic dominance that allow the comparison of more than two sets of cumulative distribution functions.

5.1 Measures of comparison of IF-sets

In this section we are going to introduce some comparison measures for IF-sets. We begin by recalling the most common comparison measures for IF-sets: distances and dissimilarities.

Definition 5.1 *A map $d : IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ is a distance between IF-sets if it satisfies the following properties:*

Positivity:	$d(A, B) \geq 0$ *for every* A, B $IFSs(\Omega)$.
Identity of indiscernibles:	$d(A, B) = 0$ *if and only if* $A = B$.
Symmetry:	$d(A, B) = d(B, A)$ *for every* A *and* B *in* $IFSs(\Omega)$.
Triangle inequality:	$d(A, C) \leq d(A, B) + d(B, C)$ *for every* A, B, C $IFSs(\Omega)$.

Definition 5.2 *A map $D: IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ is a dissimilarity for IF-sets (IF-dissimilarity, for short) if it satisfies the following axioms:*

IF-Diss.1: $D(A,A) = 0$ *for every* A $IFSs(\Omega)$.
IF-Diss.2: $D(A, B) = D(B, A)$ *for every* A, B $IFSs(\Omega)$.
IF-Diss.3: *For every* A, B, C $IFSs(\Omega)$ *such that* A B C *it holds that* $D(A, C) \geq \max(D(A, B), D(B, C))$

Remark 5.3 *Someauthors (see forinstance[93,113,211]) replace axiomIF-Diss.1by astronger condition:*

IF-Diss.1: $D(A, B) = 0 \quad A = B.$

Thus, an IF-dissimilaritythatsatisfies IF-Diss.1is more restrictivethanIF-dissimilarities. Here, we shall restrict ourselves tothe usual definition of IF-dissimilarity because it is more common in the literature.

There are seve ralexamples of dissimilarities in the literature, as we shall see in Section 5.1.3. However, since its definition is not to o restrictive, it is possible to definea counterintuitive meas ure of comparison for which axioms IF-Diss.1, IF-Diss.2 andIF-Diss.3 hold. In order to overcome this problem, we prop ose a measure of comparison of IF-sets called IF-divergence that satisfies the following natural prop erties:

- The divergence between two IF-sets is positive.
- The divergence between an IF-set and itself must be zero.
- The divergence between two IF-sets A and B is thesame than thedivergence between B and A. That is, it must be a symmetric function.
- The "more similar" two IF-sets are, the smaller is the divergence between them.

This is formally defined as foll ows.

Definition 5.4 *Let us consider a function* $D_{\text{IFS}} : IF\,Ss(\Omega) \times IF\,Ss(\Omega) \to \mathbb{R}$. *It isa divergence for IF-sets (IF-divergence for short) when it satisfies the fol lowing axioms:*

IF-Diss.1:	$D_{\text{IFS}}(A, A) = 0$	*for every* $A \quad IF\,Ss(\Omega)$.
IF-Diss.2:	$D_{\text{IFS}}(A, B) = D_{\text{IFS}}(B, A)$	*for every* $A, B \quad IF\,Ss(\Omega)$.
IF-Div.3:	$D_{\text{IFS}}(A \cap C, B \cap C) \leq D_{\text{IFS}}(A, B)$	*, for every* $A, B, C \quad IF\,Ss(\Omega)$.
IF-Div.4:	$D_{\text{IFS}}(A \quad C, B \quad C) \leq D_{\text{IFS}}(A, B)$	*, for every* $A, B, C \quad IF\,Ss(\Omega)$.

Note that IF-divergences are morerestrictive thanIF-dissimilarities. In orde r to prove this, let usfirst give a preliminary result.

Lemma 5.5 *Let* D_{IFS} *beanIF-divergence, and let* A, B, C *and* D *be IF-setssuch that* $A \quad C \quad D \quad B$. *Then* $D_{\text{IFS}}(A, B) \geq D_{\text{IFS}}(C, D)$.

Pro of Note that, if N and M are two IF-sets such that $N \quad M$, then $N \quad M = M$ and $N \cap M = N$. Then, itholdsthat:

$$C \quad D \quad C \cap D = C, \qquad D \quad B \quad B \cap D = D,$$
$$A \quad C \quad C \quad A = C, \qquad C \quad B \quad B \quad C = B.$$

Using axioms IF-Div.3 and IF-Div.4weobtain that:

$$D_{\text{IFS}}(C, D) = D_{\text{IFS}}(C \cap D, B \cap D) \leq D_{\text{IFS}}(C, B)$$
$$= D_{\text{IFS}}(A \quad C, B \quad C) \leq D_{\text{IFS}}(A, B).$$

We conclude that $D_{\text{IFS}}(C, D) \leq D_{\text{IFS}}(A, B)$. ∎

Using this le mma we can prove now that every IF-divergence is also an IF-dissimilarity.

Prop osition 5.6 *Every IF-divergence is anIF-dissimilarity.*

Pro of Let D_{IFS} be an IF-divergence, and let us check that it is also an IF-dissimilarity. For this, it suffices to prove that it satisfies axiom IF-Diss .3, because first and second axioms of IF-divergences and IF-dissimilarities coin cide. Let A, B and C be three IF-sets such that $A \quad B \quad C$. Then, takingintoaccountthat $A \quad A \quad B \quad C$, and applying the previous lemma, $D_{\text{IFS}}(A, C) \geq D_{\text{IFS}}(A, B)$. Ontheotherhand, since $A \quad B \quad C \quad C$, the previous lemma also implies that $D_{\text{IFS}}(A, C) \geq D_{\text{IFS}}(B, C)$.

Hence, D_{IFS} satisfies axiom Diss.3 and, consequently, it is a dissi milarity. ∎

We have seen that every IF-divergence is also an IF-dis similarity. In Example 5.8 we will see that the converse do es not hold in general.

In the fuzzy framework Cou so et al. ([44]) intro duced a measure of comparison called dissimilitude. It can be generalized to the comparison of IF-sets in the following way.

Definition 5.7 *A map $D : IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ is an IF-dissimilitude if it satisfies the fol lowing properties:*

IF-Diss.1: $D_{\text{IFS}}(A, A) = 0$ *for every* $A \quad IFSs(\Omega)$.

IF-Diss.2: $D_{\text{IFS}}(A, B) = D_{\text{IFS}}(B, A)$ *for every* $A, B \quad IFSs(\Omega)$.

IF-Diss.3: *If* $A, B, C \quad IFSs(\Omega)$ *satisfies* $A \quad B \quad C$, *then* $D_{\text{IFS}}(A, C) \geq \max(D_{\text{IFS}}(A, B), D_{\text{IFS}}(B, C))$.

IF-Div.4: $D_{\text{IFS}}(A \quad C, B \quad C) \leq D_{\text{IFS}}(A, B)$, *for every* $A, B, C \quad IFSs(\Omega)$.

This measure of comparison is stronger than IF-dissimilarities, butless restrictive than IF-divergences. Moreover, theconverseimplicationsdonotholdingeneral. Let usgive an example of an IF-dissim ilitude that is not an IF-divergence and an example of an IF-dissimilarity that isnotan IF-dissimilitude.

Example 5.8 *First of all, we are going to build a dissimilarity that is not a dissimilitude.*

Let us consider the function $D : IFSs(\Omega) \times IFSs(\Omega) \to [0, 1]$ *defined on a finite* Ω *by:*

$$D(A, B) = \left| \max_{\omega \quad \Omega}(\max(0, \mu_B(\omega) - \mu_A(\omega))) - \max_{\omega \quad \Omega}(\max(0, \mu_A(\omega) - \mu_B(\omega))) \right|.$$

Let us see that D *is an IF-dissimilarity:*

IF-Diss.1: $D(A, A) = 0$, *since* $\mu_B(\omega) - \mu_A(\omega) = 0$ *for any* $\omega \in \Omega$.

IF-Diss.2: *Obviously,* $D(A, B) = D(B, A)$.

IF-Diss.3: *Let* A, B *and* C *be three IF-sets such that* $A \subseteq B \subseteq C$. *Then, since* $\mu_A(\omega) \leq \mu_B(\omega) \leq \mu_C(\omega)$, *it holds that:*

$$D(A, B) = |\max_{\omega \in \Omega} \mu_B(\omega) - \mu_A(\omega)|,$$
$$D(B, C) = |\max_{\omega \in \Omega} \mu_C(\omega) - \mu_B(\omega)|,$$
$$D(A, C) = |\max_{\omega \in \Omega} \mu_C(\omega) - \mu_A(\omega)|.$$

Moreover,

$$\mu_C(\omega) - \mu_A(\omega) \geq \max(\mu_C(\omega) - \mu_B(\omega), \mu_B(\omega) - \mu_A(\omega)),$$

and therefore:

$$D(A, C) \geq \max(D(A, B), D(B, C)).$$

Thus, D *satisfies axiom IF-Diss.3 and therefore it is an IF-dissimilarity. Let us show that* D *is not a dissimilitude, or equivalently, that there are IF-sets* A, B *and* C *such that* $D(A \cup C, B \cup C) > D(A, B)$. *To see this, let us consider* $\Omega = \{\omega_1, \omega_2\}$ *and define the IF-sets* A *and* B *by:*

$$A = \{(\omega_1, 0.5, 0), (\omega_2, 0, 0)\}, \quad B = \{(\omega_1, 0, 0), (\omega_2, 0.6, 0)\},$$
$$C = \{(\omega_1, 0.5, 0), (\omega_2, 0.2, 0)\}.$$

It holds that:

$$A \cup C = \{(\omega_1, 0.5, 0), (\omega_2, 0.2, 0)\}.$$
$$B \cup C = \{(\omega_1, 0.5, 0), (\omega_2, 0.6, 0)\}.$$

Then:

$$D(A, B) = |0.5 - 0.6| = 0.1 \geq 0.4 = |0.2 - 0.6| = D(A \cup C, B \cup C).$$

Hence, D *does not fulfill Div.4, and therefore it is neither an IF-dissimilitude nor an IF-divergence.*

Example 5.9 *Let us give an IF-dissimilitude that is not an IF-divergence. Consider the function* D *defined by:*

$$D(A, B) = \begin{cases} 1 & \text{if } A = \emptyset \text{ or } B = \emptyset, \text{ but } A = B. \\ 0 & \text{otherwise.} \end{cases}$$

Let us see that this function is a dissimilitude:

IF-Diss.1: $D(A, A) = 0$ *by definition.*

IF-Diss.2: D *is symmetric by definition.*

IF-Diss.3: *Let A, B and C be three IF-sets suchthat A B C. Then,*

$$\mu_A(\omega) \leq \mu_B(\omega) \leq \mu_C(\omega) \text{ and } \nu_A(\omega) \geq \nu_B(\omega) \geq \nu_C(\omega)$$

for every ω Ω.

There are two cases:on the onehand, if $D(A, C) = 1$,then

$$D(A, C) = 1 \geq \max(D(A, B), D(B, C)).$$

On theother hand, $A=$ and $C=$ or $A = C$. Since A B C, in the first case $B=$ and in the second one $B = A = C$. In all cases, $D(A, C) = D(A, B) = D(B, C) = 0$.

Div.4: *Let us show that $D(A$ C, B $C) \leq D(A, B)$ for every IF -sets A, B and C. This inequality holds if $D(A, B) = 1$. Otherwise, if $D(A, B) = 0$ then $A=$ and $B=$ or $A = B$. Since A A C and B B C, in the first case we deduce that A $C=$ and B $C=$ andwe concludethat $D(A$ C, B $C) = D(A, B) = 0$. Inthe second case, $D(A$ C, B $C) = D(A$ C, A $C) = 0 = D(A, B)$.*

Thus, D is an IF-dissimilitude, but it is not an IF-divergence since it does not fulfill axiom Div.3: if we considerthe IF-sets A, B and C defined by

$$A = \{(\omega_0, 0, 1), (\omega, \mu_A(\omega), \nu_A(\omega)) \mid \omega = \omega_0\};$$
$$B = \{(\omega, \mu_B(\omega), \nu_B(\omega)) \mid \omega \quad \Omega\};$$
$$C = \{(\omega_0, 1, 0), (\omega, 0, 1) \mid \omega = \omega_0\};$$

where $\mu_B(\omega) > 0$ for every ω Ω and $\mu_A(\omega) = \mu_B(\omega)$ for every $\omega = \omega_0$, for a fixed element ω_0 of Ω; then, $A \cap C=$ but $B \cap C=$, and therefore:

$$D(A \cap C, B \cap C) = 1 > 0 = D(A, B).$$

Hence, D is anIF-dissimilitude that isnot an IF-divergence.

Wehave already studied the relationships among IF-dissimilarities, IF-divergences and IF-dissimilitudes, andwe have also mentioned somecounterexamples related to the distance. In fact, that there is not a general relationship b etwee n the notion of distance for IF-sets and thesethreemeasures ofcomparison. Toshowthat, westartwithanexample of an IF -distance that is not an IF-dissimilarity.

Example 5.10 *Let usconsider thefunction D defined by:*

$$D(A, B) = \begin{cases} 0 & \text{if } A = B, \\ \frac{1}{2} & \text{if } A - B = \text{ or } B - A = \text{ and } \mu_{A \cap B}(\omega) = 0.3 \quad \omega \quad \Omega, \\ 1 & \text{otherwise,} \end{cases}$$

wherethe IF-difference is theone of Example 2.56. Let us see that this fu nction isa distance for IF-sets.

Positivity: By definition, $D(A, B) \geq 0$ for every A, B $IFSs(\Omega)$.

Identity of indiscernibles: By definition, $D(A, B) = 0$ if and only if $A = B$.

Symmetry: D is alsosymmetric by definition.

Triangular inequality: Let us see that $D(A, C) \leq D(A, B) + D(B, C)$ holds forany A, B, C $IFSs(\Omega)$. On the onehand, if $D(A, C) = 0$, the inequality trivial ly holds. If $D(A, B) = \frac{1}{2}$, wecan assume, without lossof generality, that $A - C =$, and then, $A = C$. This impliesthat either $A = B$ or $B = C$, and consequently either $D(A, B) \geq \frac{1}{2}$ or $D(B, C) \geq \frac{1}{2}$. Therefore the inequality holds. Final ly, if $D(A, C) = 1$ and we assume that thetriangleinequality does not hold, thenwithoutloss of generalitywe can assume that $D(A, B) = 0$. In that case, $A = B$, and therefore $D(A, C) = D(B, C) = 1$, a contradiction arises. Weconclude that thetriangle inequality holds.

Thus, D is a distancefor IF-sets. However, itisnotanIF-dissimilarity, sincewe can find IF-sets A, B and C, with A B C, such that $D(A, C) < D(A, B)$: let us consider $\Omega = \{\omega\}$ and the IF-sets A, B and C defined by:

$$A = \{(\omega, 0.2, 0.4)\} \quad B = \{(\omega, 0.3, 0.2)\} \quad C = \{(\omega, 0.4, 0)\}$$

It is obvious that A B C. Moreover,it holds that:

$$D(A, C) = 1 \text{ and } D(B, C) = 0.5.$$

We conclude that D is notan IF-dissimilarity.

We have seen that IF-distances are not IF-dissimi laritie s in general.Thus, th ey cannot be, in general, IF-divergencesor IF-dissimilitudes, since in that case they would be in particular IF-dissimilarities. We next show that the converse implications do not hold either.

Example 5.11 *Letus giveanexample ofanIF-divergencethat isnot a distancebetween IF-sets. Considerthe function D defined by:*

$$D(A, B) = \max_{\omega \in \Omega}(\max(0, \mu_A(\omega) - \mu_B(\omega)))^2 + \max_{\omega \in \Omega}(\max(0, \mu_A(\omega) - \mu_B(\omega)))^2.$$

~~IF-Div.1:~~ It isobvious that $D(A, A) = 0$.

~~IF-Div.2:~~ By definition, D is alsosymmetric.

~~IF-Div.3:~~ Let us provethat $D(A, B) \geq D(A \cap C, B \cap C)$ for any A, B, C. Using the first part of Lemma A.1 in Appendix A, forany ω it holdsthat:

$$\max(0, \mu_A(\omega) - \mu_B(\omega)) \geq \max(0, \min(\mu_A(\omega), \mu_C(\omega)) - \min(\mu_B(\omega), \mu_C(\omega))).$$

It trivially fol lows that $D(A, B) \geq D(A \cap C, B \cap C)$.

<u>*IF-Div.4:*</u> *Similarly, let us prove that* $D(A, B) \geq D(A \quad C, B \quad C)$ *for any* A, B, C *. Taking intoaccount againthe first part of Example A.1 inLemma A, any* ω *satisfies the fol lowing:*

$$\max(0, \mu_A(\omega) - \mu_B(\omega)) \geq \max(0, \max(\mu_A(\omega), \mu_C(\omega)) - \max(\mu_B(\omega), \mu_C(\omega))).$$

This implies that $D(A, B) \geq D(A \quad C, B \quad C)$.

We conclude thatD is anIF-divergence. However, itdoesnotsatisfythetriangular inequality, because for the IF-setsA, B and C of $\Omega = \{\omega\}$*, defined by:*

$$A = \{(\omega, 0, 1)\} \quad B = \{(\omega, 0.4, 0)\} \text{ and } C = \{(\omega, 0.5, 0)\}$$

it holds that:

$$D(A, C) = 0.25 \leq 0.16 + 0.01 = D(A, B) + D(B, C).$$

Thus, D does notsatisfy the triangularinequality.

Since themeasure defined inthis example isan IF-divergence, it is alsoan IF-dissimilarity and an IF-dissimilitude. Then, wecanseethatnoneofthesemeasuressatisfy, ingeneral, the prop erties that define a distance.

Let us show next that an IF-dissimilitude and a dis tance is not necessarily an IF-divergence.

Example 5.12 *Let usconsider themap*

$$D : IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$$

defined by:

$$D(A, B) = \begin{cases} 0 & \text{if } A = B. \\ 1 & \text{if } A = B \text{ and either } \mu_A(\omega) = 0 \quad \omega \quad \Omega \text{ or } \mu_B(\omega) = 0 \quad \omega \quad \Omega. \\ 0.5 & \text{otherwise.} \end{cases}$$

First of al l, let us prove that D is a distancefor IF-sets.

<u>*Positivity:*</u> *By definition,* $D(A, B) \geq 0$ *for every* $A, B \quad IFSs(\Omega)$.

<u>*Identity of indiscernibles:*</u> *By definition,* $D(A, B) = 0$ *if and only if* $A = B$.

<u>*Triangular inequality:*</u> *Let us consider* $A, B, C \quad IFSs(\Omega)$*, and let us prove that* $D(A, C) \leq D(A, B) + D(B, C)$ *. If* $D(A, C) = 0$ *, obviouslythe inequality holds. If* $D(A, C) = 0.5$*, then* $A = C$ *, and therefore either* $A = B$ *, and consequently* $D(A, B) \geq 0.5$ *or* $B = C$ *, and consequently* $D(B, C) \geq 0.5$*. Then,* $D(A, B) + D(B, C) \geq 0.5 = D(A, C)$.

Otherwise, $D(A, C) = 1$. *In such a case,* $A = C$ *and we can assume that* $\mu_A(\omega) = 0$ *for every* $\omega \in \Omega$. *Then, if* $A = B$, $D(A, B) = 1$, *and if* $A = B$, *then* $D(B, C) = D(A, C) = 1$. *We conclude thus that the triangular inequality holds.*

Let us now prove that D *is also an IF-dissimilitude:*

IF-Diss.1: We have already seen that $D(A, A) = 0$.

IF-Diss.2: Obviously, D *is symmetric.*

IF-Diss.3: Consider $A, B, C \in IFSs(\Omega)$ *such that* $A \subseteq B \subseteq C$, *and let us prove that* $D(A, C) \geq \max(D(A, B), D(B, C))$. *Note that if* $D(A, C) = 0$, *then* $A = B = C$, *and therefore the inequality holds. Moreover, if* $D(A, C) = 1$ *then the inequality also holds because* $\max(D(A, B), D(B, C)) \leq 1$. *Finally, assume that* $D(A, C) = 0.5$. *In such a case* $A = C$, *and therefore either* $A = B$ *or* $B = C$, *and there is* $\omega \in \Omega$ *such that* $\mu_C(\omega) \geq \mu_A(\omega) > 0$. *Then, as* $\mu_C(\omega) \geq \mu_B(\omega) \geq \mu_A(\omega)$, $D(A, B), D(B, C) \leq 0.5$. *Thus, axiom IF-Diss.3 holds.*

IF-Div.4: Let us now consider three IF-sets A, B *and* C, *and let us prove that* $D(A \cup C, B \cup C) \leq D(A, B)$. *First of all, if* $D(A, B) = 1$, *then the previous inequality trivially holds, since* D *is bounded by 1. Moreover, if* $D(A, B) = 0$ *then* $A = B$, *and consequently applying IF-Diss.1* $D(A \cup C, B \cup C) = D(A \cup C, A \cup C) = 0$. *Finally, assume that* $D(A, B) = 0.5$. *In such a case,* $A = B$ *and there exist* $\omega_1, \omega_2 \in \Omega$ *such that* $\mu_A(\omega_1) > 0$ *and* $\mu_B(\omega_2) > 0$. *Let us note that:*

$$\mu_{A \cup C}(\omega) = \max(\mu_A(\omega), \mu_C(\omega)) \geq \mu_A(\omega) \text{ and}$$
$$\mu_{B \cup C}(\omega) = \max(\mu_B(\omega), \mu_C(\omega)) \geq \mu_B(\omega).$$

Consequently, $\mu_{A \cup C}(\omega_1) \geq \mu_A(\omega_1) > 0$ *and* $\mu_{B \cup C}(\omega_2) \geq \mu_B(\omega_2) > 0$. *Then it holds that* $D(A \cup C, B \cup C) \leq 0.5 = D(A, B)$.

Thus, D *is a distance and an IF-dissimilitude. Let us show that it is not an IF-divergence. Consider* $\Omega = \{\omega_1, \omega_2\}$ *and the IF-sets* A, B *and* C *defined by:*

$$A = \{(\omega_1, 1, 0), (\omega_2, 0, 0)\}.$$
$$B = \{(\omega_1, 1, 0), (\omega_2, 1, 0)\}.$$
$$C = \{(\omega_1, 0, 0), (\omega_2, 1, 0)\}.$$

Then:

$$A \cap C = \{(\omega_1, 0, 0), (\omega_2, 0, 0)\}.$$
$$B \cap C = \{(\omega_1, 0, 0), (\omega_2, 1, 0)\}.$$

Then, $D(A, B) = 0.5$ *and* $D(A \cap C, B \cap C) = 1$, *and therefore*

$$D(A \cap C, B \cap C) > D(A, B),$$

a contradiction with IF-Div.3. Thus D *cannot be an IF-divergence.*

To conclude this part, itonly remainsto showthat if D isan IF-dissimilarityanda distance, it isnot necessarily an IF-dissimilitude.

Example 5.13 *Consider themap*

$$D : IF\,Ss(\Omega) \times IF\,Ss(\Omega) \to \mathbb{R}$$

defined by:

$$D(A, B) = \begin{cases} 0 & \text{if } A = B. \\ 1 & \text{if } A = B \text{ and either } A = \Omega \text{ or } B = \Omega. \\ 0.5 & \text{otherwise.} \end{cases}$$

Let us prove that D is a distancefor IF-sets.

Positivity, the identity of indiscernibles and symmetry trivial ly hold. Letus prove that the triangular inequality is also satisfied. Let A, B and D bethreeIF-sets, andlet us see that $D(A, C) \leq D(A, B) + D(B, C)$.

- *If* $D(A, C) = 0$ *, the inequality trivial ly holds.*
- *If* $D(A, C) = 0.5$*, then* $A = C$ *, and thereforeeither* $A = B$ *or* $B = C$ *, and consequently* $D(A, B) + D(B, C) \geq 0.5 = D(A, C)$
- *Final ly, if* $D(A, C) = 1$ *, we can assume, without lossof generality, that* $A = \Omega$ *. Then, if* $B = A$ *,* $D(B, C) = 1$*, and therefore* $D(A, C) = 1 = D(A, B) + D(B, C)$. *Otherwise, if* $B = A$ *, then* $D(A, B) = 1$ *, and therefore*

$$D(A, C) = 1 \leq D(A, B) + D(B, C).$$

Thus, D is a distancefor IF-sets.

Let us now prove that it is also an IF-dissimilarity. Onthe onehand, properties IF-Diss.1 and IF-Diss.2 are trivial ly satisfied. Let us see that IF-Diss.3 also holds. Consider threeIF-sets A, B,C satisfying A B C, and let us prove that $D(A, C) \geq \max(D(A, B), D(B, C))$

- *If* $D(A, C) = 1$ *,obviously* $D(A, C) \geq \max(D(A, B), D(B, C))$
- *If* $D(A, C) = 0.5$ *, then* $A = C$ *andthere is* ω Ω *such that* $\mu_A(\omega) \leq \mu_B(\omega) \leq \mu_C(\omega) < 1$. *Then,* $\max(D(A, B), D(B, C)) \leq 0.5 = D(A, C)$.
- *Final ly, if* $D(A, C) = 0$ *,* $A = B = C$ *holds, andthen* $D(A, B) = D(B, C) = 0$.

Thus, D is a distance for IF-sets and an IF-dissimilarity. However, it is not an IF-dissimilitude, for it does not satisfy axiom IF-Div.4: to see this, consider the universe $\Omega = \{\omega_1, \omega_2\}$, and the IF-sets

$$A = \{(\omega_1, 1, 0), (\omega_2, 0, 0)\} \text{ and } B = \{(\omega_1, 0, 0), (\omega_2, 1, 0)\}.$$

It holds that $D(A, B) = 0.5$. However, if we consider $C = B$, then $A \quad C = \Omega$, and therefore:

$$D(A \quad C, B \quad C) = D(\Omega, B) = 1.$$

Then, $D(A \quad C, B \quad C) > D(A, B)$, and therefore axiom IF-Div.4 is not satisfied. This shows that D is not an IF-divergence.

Figure 5.1 summarizes the relationships between the different methods for comparing IF-sets.

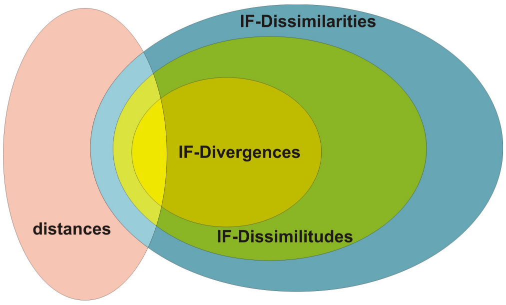


Figure 5.1: Relationships among IF-divergences, IF-dissimilitudes, IF-dissimilarities and distances for IF-sets.

5.1.1 Theoretical approach to the comparison of IF-sets

Bouchon-Meunier et al. ([27]) proposed a general measure of comparison for fuzzy sets that generates some particular measures depending on the conditions imposed to such a general measure.

Following this ideas, in this section we define a general measure of comparison between IF-sets that, depending on the imposed properties generates either distances, or IF-dissimilarities or IF-divergences.

For this, let us consider a function $D : IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$, and assume that there is a generator function G_D:

$$G_D : IFSs(\Omega) \times IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}^+ \tag{5.1}$$

such that D can be expressed by:

$$D(A,B) = G_D(A \cap B, B - A, A - B),$$

where $-$ is a difference operator for IF-sets, according to Definition 2.55, that fulfills D3, D4 and D5.

We shall see that depending on the conditions imposed on G_D, we can obtain that D is either an IF-dissimilarity, an IF-divergence or a distance for IF-sets.

We begin by determining which conditions must be imposed on G_D in order to obtain a distance for IF-sets.

Proposition 5.14 *Consider the function $D: IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ that can be expressed as in Equation (5.1) by means of a generator $G_D : IFSs(\Omega) \times IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}^+$. If the function G_D satisfies the properties:*

S-Dist.1: $G_D(A,B,C) = 0$ *if and only if* $B = C =$;

S-Dist.2: $G_D(A,B,C) = G_D(A,C,B)$ *for every* A,B,C $IFSs(\Omega)$;

S-Dist.3: *For every* A,B,C $IFSs(\Omega)$,

$$G_D(A \cap C, C - A, A - C) \leq G_D(A \cap B, B - A, A - B) + G_D(B \cap C, C - B, B - C);$$

then D is a distance for IF-sets.

Proof Let us prove that D satisfies the axioms of IF-distances.

Positivity: it trivially follows from the positivity of G_D. To show the identity of indiscernibles, let A and B be two IF-sets. Then, by property S-Dist.1:

$$D(A,B) = G_D(A \cap B, B - A, A - B) = 0 \quad B - A = A - B = \ ,$$

and by properties D1 and D5 this is equivalent to $A = B$.

Symmetry: Let A and B be two IF-sets. Using S-Dist.2, we have that:

$$\begin{aligned} D(A,B) &= G_D(A \cap B, B - A, A - B) \\ &= G_D(A \cap B, A - B, B - A) = D(B,A). \end{aligned}$$

Triangular inequality: Let A, B and C be three IF-sets. By S-Dist.3, it holds that:

$$\begin{aligned} D(A,C) &= G_D(A \cap C, C - A, A - C) \\ &\leq G_D(A \cap B, B - A, A - B) + G_D(B \cap C, C - B, B - C) \\ &= D(A,B) + D(B,C). \end{aligned}$$

Let us now consider IF-dissimilarities. We have proven the following result:

Prop osition 5.15 *Let D be amap $D : IF\,Ss(\Omega) \times IF\,Ss(\Omega) \to \mathbb{R}^+$ that canbeexpressed asinEquation* (5.1) *bymeansofthe generator G_D, where $G_D : IF\,Ss(\Omega) \times IF\,Ss(\Omega) \times IF\,Ss(\Omega) \to \mathbb{R}^+$. Then, D is anIF-dissimilarity if G_D satisfies the fol lowing properties:*

S-Diss.1: $G_D(A, \ , \ ) = 0$ *for every* $A \quad IF\,Ss(\Omega)$.
S-Dist.2: $G_D(A, B, C) = G_D(A, C, B)$ *for every* $A, B, C \quad IF\,Ss(\Omega)$.
S-Diss.3: $G_D(A, B, \ )$ *is increasing in* B.
S-Diss.4: $G_D(A, B, \ )$ *is decreasing in* A.

Pro of Letus provethat D isan IF-dissimilarity.

~~IF-Diss.1:~~ Let A be an IF-set. By D1 and S-Diss .1 it holds that

$$D(A, A) = G_D(A \cap A, A - A, A - A) = G_D(A, \ , \ ) = 0.$$

~~IF-Diss.2:~~ Let A and B be two IF-sets. Then, byS-Dist.2, D is symmetric:

$$\begin{aligned} D(A, B) &= G_D(A \cap B, B - A, A - B) \\ &= G_D(A \cap B, A - B, B - A) = D(B, A). \end{aligned}$$

~~IF-Diss.3:~~ Let A, B and C b e three I F-s ets such that $A \quad B \quad C$, and let us prove that $D(A, C) \geq \max(D(A, B), D(B, C))$. Firstof all, letus compute $D(A, C)$, $D(A, B)$ and $D(B, C)$.

$$\begin{aligned} D(A, C) &= G_D(A \cap C, C - A, A - C) = G_D(A, C - A, \ ). \\ D(A, B) &= G_D(A \cap B, B - A, A - B) = G_D(A, B - A, \ ). \\ D(B, C) &= G_D(B \cap C, C - B, B - C) = G_D(B, C - B, \ ). \end{aligned}$$

Ononehand, letus provethat $D(A, C) \geq D(A, B)$. ByD2, itholdsthat $B - A \quad C - A$, and therefore, byS-Diss.3:

$$D(A, C) = G_D(A, C - A, \ ) \geq G_D(A, B - A, \ ) = D(A, B).$$

Let us prove next that $D(A, C) \geq D(B, C)$. ByD4it holdsthat $C - B \quad C - A$, and therefore:

$$\begin{aligned} D(A, C) = G_D(A, C - A, \ ) &\overset{S-Diss.4}{\geq} G_D(B, C - A, \ ) \\ &\overset{S-Diss.3}{\geq} G_D(B, C - B, \ ) = D(B, C). \end{aligned}$$

Thus, we conclude that D isan IF-dissimilarity. ∎

ConcerningIF-divergences, we haveestablished thefollowing:

Prop osition 5.16 *Let D be a map $D : IF\,Ss(\Omega) \times IF\,Ss(\Omega) \to \mathbb{R}$ generated byG_D as in Equation* (5.1), *where $G_D : IF\,Ss(\Omega) \times IF\,Ss(\Omega) \times IF\,Ss(\Omega) \to \mathbb{R}^+$. Then, D is an IF-divergence if G_D satisfies the fol lowing properties:*

S-Diss.1: $G_D(A, \ , \ ) = 0$ *for every* A, B $IF\,Ss(\Omega)$.
S-Dist.2: $G_D(A, B, C) = G_D(A, C, B)$ *for every* A, B, C $IF\,Ss(\Omega)$.
S-Div.3: $G_D(A, B, C)$ *is increasing in* B *and* C.
S-Div.4: $G_D(A, B, C)$ *is independent of* A.

Note that axiom S-Div.4 is a very strongcondition. We require it because IF-divergences fo cus on the difference b etween the IF-sets instead of the intersection.

Pro of Let us prove that D isan IF-divergence.

First andsecond axiomsof IF-divergences and IF-dissimilarities coincide. Furthermore, as we proved in Prop osition 5.15, they follow from S-Diss.1 and S-Dist.2.

~~IF-Div.3:~~ Let A, B and C be three IF-sets. Since the IF-difference op erator fulfills D3, then $(A \cap C) - (B \cap C)$ $A - B$ and $(B \cap C) - (A \cap C)$ $B - A$. Therefore, by S-Div.3 and S-Div.4:

$$
\begin{aligned}
& D(A \cap C, B \cap C) \\
& \quad = G_D(A \cap B \cap C, (B \cap C) - (A \cap C), (A \cap C) - (B \cap C)) \\
& \quad = G_D(A \cap B, (B \cap C) - (A \cap C), (A \cap C) - (B \cap C)) \\
& \quad \leq G_D(A \cap B, B - A, A - B) = D(A, B).
\end{aligned}
$$

~~IF-Div.4:~~ Consider the IF-sets A, B and C. Asinthe previousaxiom, applying prop erty D4 of the IF-difference $-$, we obtainthat $(A \quad C) - (B \quad C)$ $A - B$ and $(B \quad C) - (A \quad C)$ $B - A$. As a consequence,

$$
\begin{aligned}
& D(A \quad C, B \quad C) \\
& \quad = G_D((A \quad C) \cap (B \quad C), (B \quad C) - (A \quad C), (A \quad C) - (B \quad C)) \\
& \quad \stackrel{S-Div.4}{=} G_D(A \cap B, (B \quad C) - (A \quad C), (A \quad C) - (B \quad C)) \\
& \quad \stackrel{S-Div.3}{\leq} G_D(A \cap B, B - A, A - B) = D(A, B).
\end{aligned}
$$

We conclude that D isan IF-divergence. ∎

Inorder to findsufficient conditionsover G_D so as to build an IF-d issimilitude D, we need D tosatisfy axiomsIF-Diss.1, IF-Diss.2, IF-Diss.3and IF-Div.4. Aswe have already mention ed, axioms IF-Diss.1 and IF-Diss.2 are implied by conditions:

S-Diss.1: $G_D(A, \ , \ )$ for every A, B $IF\,Ss(\Omega)$.
S-Dist.2: $G_D(A, B, C) = G_D(A, C, B)$ for every A, B $IF\,Ss(\Omega)$.

In order to prove condition IF-Div.4, in Prop osition 5.16 we required the following:

S-Div.3: $G_D(A, B, C)$ is increasing in B and C.
S-Div.4: $G_D(A, B, C)$ is indep endent of A.

Moreover, itis trivial that these conditions implyS-Diss.3 and S-Diss.4, that also follow from axiom IF-Diss.3. Therefore, the conditions that need to be imp osed on G_D

in order to obtain an IF-dissimilitude are the same that we have imp osed in order to obtain an IF -divergence.

Letusgive an example of a function G_D thatgeneratesan IF-dissimilaritybutnot an IF-divergenc e.

Example 5.17 *Consider thefunction* $G_D : IF\,Ss(\Omega) \times IF\,Ss(\Omega) \times IF\,Ss(\Omega) \to \mathbb{R}^+$ *defined, for every* A, B, C $IF\,Ss(\Omega)$*, by:*

$$G_D(A, B, C) = \left| \max_{\omega \ \Omega} \mu_B(\omega) - \max_{\omega \ \Omega} \mu_C(\omega) \right|.$$

This function generates an IF-dissimilarity because it satisfies properties S-Diss. i, with $i = 1, 3, 4$ *and S-Dist.2.*

S-Diss.1: By definition, $G_D(A, \ , \ ) = 0$*, since* $\mu \ (\omega) = 0$ *for every* $\omega \ \Omega$.

S-Dist.2: G_D *is symmetric with respect its second and third component s:*

$$\begin{aligned} G_D(A, B, C) &= \left| \max_{\omega \ \Omega} \mu_B(\omega) - \max_{\omega \ \Omega} \mu_C(\omega) \right| \\ &= \left| \max_{\omega \ \Omega} \mu_C(\omega) - \max_{\omega \ \Omega} \mu_B(\omega) \right| = G_D(A, C, B). \end{aligned}$$

S-Diss.3: Let A, B *and* B *be threeIF-sets such that* $B \quad B$*. Then,* $\mu_B(\omega) \leq \mu_B(\omega)$ *for every* $\omega \ \Omega$*. Then itholds that:*

$$G_D(A, B, \ ) = \max_{\omega \ \Omega} \mu_B(\omega) \leq \max_{\omega \ \Omega} \mu_B(\omega) = G_D(A, B, \ ).$$

Thus, $G_D(A, B, \ )$ *is increasing in* B*.*

S-Diss.4: It is obvious that G_D *doesnot dependon itsfirst component, and therefore, it is in particular decreasing on* A*.*

Hence, G_D *satisfiesthe conditions of Proposition 5.15, and therefore themap* D *defined by:*

$$D(A, B) = G_D(A \cap B, B - A, A - B), \quad \text{for every } A, B \quad IF\,Ss(\Omega)$$

is an IF-dissimilarity. However, ingeneral G_D *doesnotsatisfy S-Div.4. Toseethis, it is enough to consider the IF-difference of Example 2.56. In that case, the function* G_D *generatesthe IF-dissimilarity of Example 5.8, which was showed not to satisfy condition IF-Div.4. Then,* D *isneitheranIF-dissimilitudenoran IF-divergence. This implies that* G_D *does not fulfill S-Div.4, because otherwise* D *would bean IF-divergence.*

Let us see next an example of afunction G_D that generates an IF-divergence th at is not adistance for IF-sets.

Example 5.18 *Consider the function* $G_D : IF\,Ss(\Omega) \times IF\,Ss(\Omega) \times IF\,Ss(\Omega) \rightarrow R^+$ *defined by:*

$$G_D(A, B, C) = \max_{\omega\ \ \Omega} \mu_B(\omega)^2 + \max_{\omega\ \ \Omega} \mu_C(\omega)^2,$$

for every $A, B, C \quad IF\,Ss(\Omega)$. *This function generates an IF-divergence, since it trivial ly satisfies the conditions in Proposition 5.16. However, it does not generat e a distance for IF-sets. Tosee it, consider theIF-differencedefined in Example 2.56. Then, the IF-divergence that generates*G_D *with this IF-differencecoincides with the one given in Example 5.11, where weproved that it was not adistance for IF-sets.*

Finally, letus give an exampleof a fun ction G_D that generatesa distance forfuzzy setsthat is notan IF-dissimilarity, andthereforeit isneitheran IF-divergencenoran IF-dissimilitude.

Example 5.19 *Consider thefunction*

$$G_D : IF\,Ss(\Omega) \times IF\,Ss(\Omega) \times IF\,Ss(\Omega) \rightarrow R^+$$

by:

$$G_D(A, B, C) = \begin{cases} 0 & \text{if } B = C = \quad , \\ 0.5 & \text{if } B = \quad \text{or } C = \quad \text{and } \mu_A(\omega) = 0.3 \text{ for all } \omega \quad \Omega, \\ 1 & \text{otherwise.} \end{cases}$$

Let us prove that G_D *satisfies conditionsof Proposition5.14.*

~~S-Dist.1~~: *By definition,* $G_D(A, B, C) = 0$ *if and only if* $B = C = \quad$.

~~S-Dist.2~~: *Obviously,* $G_D(A, B, C) = G_D(A, C, B)$ *for every* $A, B, C \quad IF\,Ss(\Omega)$.

~~S-Dist.3~~: *Let us consider* $A, B, C \quad IF\,Ss(\Omega)$, *and we want to prove that*

$$G_D(A \cap C, C - A, A - C) \leq G_D(A \cap B, B - A, A - B) + G_D(C \cap B, B - C, C - B).$$

- *If* $G_D(A \cap C, C - A, A - C) = 0$ *, then the inequality trivial ly holds.*
- *Let us nowassume that* $G_D(A \cap C, C - A, A - C) = 0.5$. *Thus, either* $A - C =$ *or* $C - A =$ *and* $\mu_{A \cap C}(\omega) = 0.3$ *for every* $\omega \quad \Omega$. *Let us note t hat, as* $A = C$ *, either* $A = B$ *or* $B = C$. *Equivalently, either* $G_D(A \cap B, B - A, A - B) \geq 0.5$ *or* $G_D(C \cap B, B - C, C - B) \geq 0.5$. *Then, inthis casethe inequality alsoholds.*
- *Final ly, consider the case where* $G_D(A \cap C, C - A, A - C) = 1$. *Then,* $A - C =$ *or* $C - A =$ *and* $\mu_{A \cap C}(\omega) = 0.3$ *for some* $\omega \quad \Omega$. *If* $A = B$ *, then:*

$$G_D(A \cap B, B - A, A - B) = 0 \quad \text{and}$$
$$G_D(C \cap B, B - C, C - B) = G_D(C \cap A, A - C, C - A) = 1.$$

> *The samehappens when* $B = C$. *Otherwise, if* $A = B$ *and* $B = C$ *, then both* $G_D(C \cap B, B - C, C - B)$ *and* $G_D(A \cap B, B - A, A - B)$ *aregreater or equal to* 0.5*, and its sum equals 1.*

Therefore, G_D *generatesa distance forIF-sets. Toshow that it generates neither an IF-dissimilaritynor an IF-divergence, itis enoughto considertheIF-differenceof Example2.56, becauseinthat casethefunction* G_D *generatesthe distance ofExamples5.10, wherewe showed that such function is neither an IF-dissimilarity nor an IF-divergence.*

We have see n sufficient conditions for G_D to generatedistances, IF-dissimi laritie s and IF-divergences. However, such conditions are not necess ary, and we c an not assure that every distance , IF-dissimilarity or IF-divergence can b e generated in this way.

Aswe haveseen, IF-divergencesare morerestrictive thanIF-dissimilaritiesand IF-dissimilitudes. Thus, IF-divergencesavoidsomecounterintuitivemeasuresofcomparison ofIF-sets, sincethe strongertheconditions, themore"robust" themeasureis. Because of this, we think it is preferable to work with IF-divergences, and we shall fo cus on them in the remainder of this chapter.

5.1.2 Properties of the IF-divergences

We have prop osed an axiomatic definition of divergence measures for intuitionistic fuzzy sets, which are particular cas es of dissimi larity and dissimilitude measures. Next, we study their prop erties in more detail. We begin by noting that a desirable prop erty fora measure of the difference between IF-sets is positivity. Although it has not b een imp osed in the definition, it can be easily derived from axioms IF-Diss.1 and IF-Div.3:

Lemma 5.20 *If* $D : IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ *satisfiesIF-Diss.1 and IF-Div.3, thenit is posit ive.*

Pro of Consider two IF-sets A and B. From IF-Div.3, for every C $IFSs(\Omega)$ it holds that:

$$D(A, B) \geq D(A \cap C, B \cap C).$$

If we take $C =$, then:

$$D(A, B) \geq D(A \cap \ , B \cap \ ) = D(\ , \ ) = 0,$$

by IF-Diss .1. Thus, D is a positive function. ∎

Now we investigate an interesting prop erty of IF-divergences.

Prop osition 5.21 *Given anIF-divergence* D_{IFS}*, it fulfil ls that:*

$$D_{IFS}(A \cap B, B) = D_{IFS}(A, A \quad B),$$

and this value is lower than or equal to $D_{IFS}(A,B)$ *and* $D_{IFS}(A\cap B, A\cup B)$*, that is:*

$$D_{IFS}(A\cap B,B)=D_{IFS}(A,A\cup B)\leq \min\{D_{IFS}(A,B), D_{IFS}(A\cap B,A\cup B)\}.$$

However, there is no fixed relationship between $D_{IFS}(A\cap B,A\cup B)$ *and* $D_{IFS}(A,B)$.

Proof By the definitions of union and intersection of intuitionistic fuzzy sets, we have that $(A\cup B)\cap B=B$ and $(A\cap B)\cup A=A$. Applying axioms IF-Div.3 and IF-Div.4, we obtain that

$$\begin{aligned} D_{IFS}(A\cap B,B) &= D_{IFS}(A\cap B,(A\cup B)\cap B)\leq D_{IFS}(A,A\cup B)\\ &= D_{IFS}((A\cap B)\cup A, B\cup A)\leq D_{IFS}(A\cap B,B). \end{aligned}$$

Thus, $D_{IFS}(A\cap B,B)=D_{IFS}(A,A\cup B)$.

On the other hand, $B\cap B=B$, whence

$$D_{IFS}(A\cap B,B)=D_{IFS}(A\cap B,B\cap B)\leq D_{IFS}(A,B) \text{ by Axiom IF-Div.3}.$$

Finally, since $A\cap B\subseteq A\subseteq A\cup B$, by Lemma 5.5 we have that

$$D_{IFS}(A,A\cup B)\leq D_{IFS}(A\cap B,A\cup B).$$

In order to prove that there is no dominance relationship between $D_{IFS}(A\cap B,A\cup B)$ and $D_{IFS}(A,B)$, let us consider the universe $\Omega=\{\omega\}$ and the IF-sets:

$$\begin{aligned} A &= \{(\omega,0.2,0.6)\} & A\cap B &= \{(\omega,0.2,0.7)\}\\ B &= \{(\omega,0.3,0.7)\} & A\cup B &= \{(\omega,0.3,0.6)\} \end{aligned}$$

Consider the IF-divergences D_L and l_{IFS} defined by:

$$\begin{aligned} D_L(A,B) &= \tfrac{1}{4}\big(|(\mu_A(\omega)-\nu_A(\omega))-(\mu_B(\omega)-\nu_B(\omega))|+|\mu_A(\omega)-\mu_B(\omega)|\\ &\quad +|\nu_A(\omega)-\nu_B(\omega)|\big).\\ l_{IFS}(A,B) &= \tfrac{1}{2}|\mu_A(\omega)-\mu_B(\omega)|+|\nu_A(\omega)-\nu_B(\omega)|+|\pi_A(\omega)-\pi_B(\omega)|. \end{aligned}$$

As we shall see in Section 5.1.3, they correspond to the Hong and Kim IF-divergence and the Hamming distance, respectively. Then:

$$\begin{aligned} & l_{IFS}(A,B)=0.2 \text{ and}\\ & l_{IFS}(A\cap B,A\cup B)=0.1.\\ & D_L(A,B)=\tfrac{0.2}{4} \text{ and}\\ & D_L(A\cap B,A\cup B)=\tfrac{0.2+0.1+0.1}{4}=\tfrac{0.4}{4}. \end{aligned}$$

Thus:

$$l_{IFS}(A,B)>l_{IFS}(A\cap B,A\cup B) \text{ and } D_L(A,B)<D_L(A\cap B,A\cup B)$$

and therefore, there is not fixed relationship between these two quantities. ∎

Next, we shall study under which conditions axioms IF-Div.3 and IF-Div.4 are equivalent. But before tackling this problem, we give an example showing that they are not equivalent in general.

Example 5.22 *Consider the function* $D : IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ *given by*

$$D(A, B) = \sum_{\omega \in \Omega} h(\mu_A(\omega), \mu_B(\omega)), \text{ for every } A, B \in IFSs(\Omega),$$

where h *is defined by*

$$h(x, y) = \begin{cases} 0 & \text{if } x = y. \\ 1 - xy & \text{if } x \neq y. \end{cases}$$

We shall prove in Example 5.53 of Section 5.1.5 that D *satisfies IF-Diss.1, IF-Diss.2 and IF-Div.4. However, it is not an IF-divergence. For instance, if we consider a universe* $\Omega = \{\omega_1, \ldots, \omega_n\}$*, and the IF-sets defined by:*

$$A = \{(\omega, 0.2, 0.8) \mid \omega \in \Omega\}.$$
$$B = \{(\omega, 0.8, 0.2) \mid \omega \in \Omega\}.$$
$$C = \{(\omega, 0.5, 0.5) \mid \omega \in \Omega\}$$

it holds that:

$$D(A \cap C, B \cap C) = D(A, C) = \sum_{\omega \in \Omega} (1 - 0.2 \cdot 0.5) = \sum_{\omega \in \Omega} 0.9 = 0.9n.$$

$$D_{IF}(A, B) = \sum_{\omega \in \Omega} (1 - 0.2 \cdot 0.8) = \sum_{\omega \in \Omega} 0.84 = 0.84n.$$

Thus, $D(A \cap C, B \cap C) = 0.9n > 0.84n = D(A, B)$ *and therefore IF-Div.3 is not satisfied.*

Hence, we have an example of a function that satisfies IF-Div.4 but it does not satisfy IF-Div.3. Next we are going to show by means of an example that IF-Div.3 does not imply IF-Div.4 either. Consider the function $D : IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ *given by:*

$$D(A, B) = \sum_{\omega \in \Omega} h(\mu_A(\omega), \mu_B(\omega)) \text{ for every } A, B \in FS(\Omega),$$

where $h : \mathbb{R}^2 \to \mathbb{R}$ *is defined by:*

$$h(x, y) = \begin{cases} 0 & \text{if } x = y. \\ xy & \text{if } x \neq y. \end{cases}$$

We shall also see in Example 5.53 of Section 5.1.5 that this function satisfies IF-Diss.1, IF-Diss.2 and IF-Div.3, but it is not an IF-divergence: consider $\Omega = \{\omega_1, \ldots, \omega_n\}$*, and the IF-sets of the previous example. Then, it holds that*

$$D(A \cup C, B \cup C) = D(C, B) = \sum_{\omega \in \Omega} 0.8 \cdot 0.5 = \sum_{\omega \in \Omega} 0.4 = 0.4n.$$

$$D(A, B) = \sum_{\omega \in \Omega} 0.2 \cdot 0.8 = \sum_{\omega \in \Omega} 0.16 = 0.16n.$$

We can conclude that axiom IF-Div.4 is not satisfied since

$$D(A \cup C, B \cup C) = 0.4n > 0.16n = D(A, B).$$

Therefore, axioms IF-Div.3 and IF-Div.4 are not related in general. We shall see however, that under some additional conditions they become equivalent. Let us consider the following natural property:

IF-Div.5: $D_{IFS}(A, B) = D_{IFS}(A^c, B^c)$ for every A, B $IFSs(\Omega)$.

In the following section we shall see some examples of IF-divergences satisfying this property. To see, however, that not all IF-divergences satisfy IF-Div.5, take $\Omega = \{\omega\}$ and the function defined by:

$$D_{IFS}(A, B) = |\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)|^2. \qquad (5.2)$$

We shall prove in Example 5.54 of Section 5.1.5 that this function is an IF-divergence. However, it does not satisfy IF-Div.5. To see that, consider the IF-sets

$$A = \{(\omega, 0.6, 0.4)\} \text{ and } B = \{(\omega, 0.5, 0.1)\}$$

It holds that:

$$D_{IFS}(A, B) = 0.1 + 0.09 = 0.19 = 0.31 = 0.3 + 0.01 \neq D_{IFS}(A^c, B^c).$$

Our next result shows that, when IF-Div.5 is satisfied, then axioms IF-Div.3 and IF-Div.4 are equivalent.

Proposition 5.23 *If D is a function $D: IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ satisfying the property IF-Div.5, then it satisfies IF-Div.3 if and only if it satisfies IF-Div.4.*

Proof First of all let us show that, since $D(A, B) = D(A^c, B^c)$ by IF-Div.5, it also holds that:

$$D(A \quad C, B \quad C) = D((A \quad C)^c, (B \quad C)^c) = D(A^c \cap C^c, B^c \cap C^c).$$

Assume that D satisfies IF-Div.3:

$$D(A \cap C, B \cap C) \leq D(A, B) \text{ for every } A, B \quad IFSs(\Omega).$$

Then it also satisfies IF-Div.4:

$$D(A \quad C, B \quad C) = D(A^c \cap C^c, B^c \cap C^c) \leq D(A^c, B^c) = D(A, B).$$

Similarly, assume that D satisfies IF-Div.4, that is,

$$D(A \quad C, B \quad C) \leq D(A, B) \text{ for every } A, B \quad IFSs(\Omega).$$

Then, it also satisfies axiom IF-Div.3:

$$D(A \cap C, B \cap C) = D(A^c \quad C^c, B^c \quad C^c) \leq D(A^c, B^c) = D(A, B).$$

Now, we will obtain a general expression of IF-divergences by comparing the membership and non-membership functions of the IF-sets by means of a t-conorm.

Proposition 5.24 *Consider a finite set Ω. If S and S are two t-conorms, the function D_{IFS} defined by:*

$$D_{\text{IFS}}(A,B) = S_{\omega \in \Omega}(S(|\mu_A(\omega) - \mu_B(\omega)|, |\nu_A(\omega) - \nu_B(\omega)|))$$

for every $A, B \in IFSs(\Omega)$, is an IF-divergence. Moreover, it satisfies IF-Div.5.

Proof Let us prove that D_{IFS} fulfills axioms IF-Diss.1 to IF-Div.4.

~~IF-Diss.1:~~ Let A be an IF-set. Obviously, $D_{\text{IFS}}(A,A) = 0$:

$$D_{\text{IFS}}(A,A) = S_{\omega \in \Omega}(S(0,0)) = S(0,\ldots,0) = 0.$$

~~IF-Diss.2:~~ Let A and B be two IF-sets. It holds that:

$$\begin{aligned} D_{\text{IFS}}(A,B) &= S_{\omega \in \Omega}(S(|\mu_A(\omega) - \mu_B(\omega)|, |\nu_A(\omega) - \nu_B(\omega)|)) \\ &= S_{\omega \in \Omega}(S(|\mu_B(\omega) - \mu_A(\omega)|, |\nu_B(\omega) - \nu_A(\omega)|)) = D_{\text{IFS}}(B,A). \end{aligned}$$

~~IF-Div.3:~~ Let A, B and C three IF-sets. We have to prove that

$$D_{\text{IFS}}(A,B) \geq D_{\text{IFS}}(A \cap C, B \cap C).$$

Applying the first part of Lemma A.1 of Appendix A, we have that

$$\begin{aligned} |\mu_A(\omega) - \mu_B(\omega)| &\geq |\min(\mu_A(\omega), \mu_C(\omega)) - \min(\mu_B(\omega), \mu_C(\omega))| = |\mu_{A \cap C}(\omega) - \mu_{B \cap C}(\omega)|. \\ |\nu_A(\omega) - \nu_B(\omega)| &\geq |\max(\nu_A(\omega), \nu_C(\omega)) - \max(\nu_B(\omega), \nu_C(\omega))| = |\nu_{A \cap C}(\omega) - \nu_{B \cap C}(\omega)|. \end{aligned}$$

Since every t-conorm is increasing, it holds that:

$$\begin{aligned} D_{\text{IFS}}(A,B) &= S_{\omega \in \Omega}(S(|\mu_A(\omega) - \mu_B(\omega)|, |\nu_A(\omega) - \nu_B(\omega)|)) \\ &\geq S_{\omega \in \Omega}(S(|\mu_{A \cap C}(\omega) - \mu_{B \cap C}(\omega)|, |\nu_{A \cap C}(\omega) - \nu_{B \cap C}(\omega)|)) \\ &= D_{\text{IFS}}(A \cap C, B \cap C). \end{aligned}$$

~~IF-Div.4:~~ Consider three IF-sets A, B and C. Using the first part of Lemma A.1 of Appendix A, we see that:

$$\begin{aligned} |\mu_A(\omega) - \mu_B(\omega)| &\geq |\max(\mu_A(\omega), \mu_C(\omega)) - \max(\mu_B(\omega), \mu_C(\omega))| \\ &= |\mu_{A \cup C}(\omega) - \mu_{B \cup C}(\omega)|. \\ |\nu_A(\omega) - \nu_B(\omega)| &\geq |\min(\nu_A(\omega), \nu_C(\omega)) - \min(\nu_B(\omega), \nu_C(\omega))| \\ &= |\nu_{A \cup C}(\omega) - \nu_{B \cup C}(\omega)|. \end{aligned}$$

Since t-conorms are increasing operators,

$$\begin{aligned} D_{\text{IFS}}(A,B) &= S_{\omega \in \Omega}(S(|\mu_A(\omega) - \mu_B(\omega)|, |\nu_A(\omega) - \nu_B(\omega)|)) \\ &\geq S_{\omega \in \Omega}(S(|\mu_{A \cup C}(\omega) - \mu_{B \cup C}(\omega)|, |\nu_{A \cup C}(\omega) - \nu_{B \cup C}(\omega)|)) \\ &= D_{\text{IFS}}(A \cup C, B \cup C). \end{aligned}$$

Thus, D_{IFS} is an IF-divergence. Now, we are going to prove that it also satisfies IF-Div.5. Using that every t-conorm is symmetric, we deduce that:

$$\begin{aligned} D_{IFS}(A,B) &= S_{\omega\in\Omega}(S(|\mu_A(\omega)-\mu_B(\omega)|,|\nu_A(\omega)-\nu_B(\omega)|)) \\ &= S_{\omega\in\Omega}(S(|\nu_A(\omega)-\nu_B(\omega)|,|\mu_A(\omega)-\mu_B(\omega)|)) = D_{IFS}(A^c,B^c). \end{aligned}$$

Therefore $D_{IFS}(A,B) = D_{IFS}(A^c,B^c)$ for every $A,B \in IFSs(\Omega)$. ∎

One of the conditions we required on IF-divergences was that "the more similar two IF-sets are, the lower the divergence is between them". In the following result we are going to see that, if the non-membership functions of A and B are the same than the ones of C and D, respectively, or the membership functions of C and D are the same, then the IF-divergence between A and B is greater than the IF-divergence between C and D.

Proposition 5.25 *Let A and B be two IF-sets. Let us consider the IF-sets C_A and D_B given by:*

$$C_A = \{(\omega,\mu(\omega),\nu_A(\omega)) \mid \omega\in\Omega\},$$
$$D_B = \{(\omega,\mu(\omega),\nu_B(\omega)) \mid \omega\in\Omega\},$$

where $\mu:\Omega\to[0,1]$ is a map such that $\mu(\omega)+\nu_A(\omega)\le 1$ and $\mu(\omega)+\nu_B(\omega)\le 1$ for every $\omega\in\Omega$. If D is an IF-divergence, then $D(A,B)\ge D(C_A,D_B)$.

Proof Let us define the following IF-set:

$$N = \{(\omega,\min(\mu_A(\omega),\mu(\omega)),0) \mid \omega\in\Omega\}.$$

Then,

$$A\cap N = \{(\omega,\min(\mu_A(\omega),\mu(\omega)),\nu_A(\omega)) \mid \omega\in\Omega\}.$$
$$B\cap N = \{(\omega,\min(\mu_B(\omega),\mu(\omega),\mu_A(\omega)),\nu_B(\omega)) \mid \omega\in\Omega\}.$$

Applying IF-Div.3 we obtain that $D(A,B)\ge D(A\cap N,B\cap N)$. Consider now another IF-set, defined by:

$$M = \{(\omega,\mu(\omega),\max(\nu_A(\omega),\nu_B(\omega))) \mid \omega\in\Omega\}.$$

We obtain that:

$$\begin{aligned} (A\cap N)\cup M &= \{(\omega,\max(\mu(\omega),\min(\mu_A(\omega),\mu(\omega))),\nu_A(\omega)) \mid \omega\in\Omega\} \\ &= \{(\omega,\mu(\omega),\nu_A(\omega)) \mid \omega\in\Omega\} = C_A. \\ (B\cap N)\cup M &= \{(\omega,\max(\mu(\omega),\min(\mu_B(\omega),\mu(\omega),\mu_A(\omega))),\nu_B(\omega)) \mid \omega\in\Omega\} \\ &= \{(\omega,\mu(\omega),\nu_B(\omega)) \mid \omega\in\Omega\} = D_B. \end{aligned}$$

Applying IF-Div.4,

$$D(A,B)\ge D(A\cap N,B\cap N)\ge D((A\cap N)\cup M,(B\cap N)\cup M) = D(C_A,D_B).$$

Analogously, we can obtain a similar result by exchanging the membership and the non-membership functions.

Proposition 5.26 *Let A and B be two IF-sets. Let us consider the IF-sets C_A and D_B given by:*

$$C = \{(\omega, \mu_A(\omega), \nu(\omega)) \mid \omega \in \Omega\} \text{ and } D = \{(\omega, \mu_B(\omega), \nu(\omega)) \mid \omega \in \Omega\},$$

where $\nu : \Omega \to [0,1]$ is a map such that $\mu_A(\omega) + \nu(\omega) \leq 1$ and $\mu_B(\omega) + \nu(\omega) \leq 1$ for every $\omega \in \Omega$. If D_{IFS} is an IF-divergence, then $D_{IFS}(A, B) \geq D_{IFS}(C_A, D_B)$.

We conclude this section with a property that assures that some transformations of IF-divergences are also IF-divergences.

Proposition 5.27 *If D is an IF-divergence and $\varphi : \mathbb{R} \to \mathbb{R}$ is an increasing function with $\varphi(0) = 0$, then D^{φ} defined by:*

$$D^{\varphi}_{IFS}(A, B) = \varphi(D_{IFS}(A, B)) \text{ for every } A, B \in IFSs(\Omega),$$

is also an IF-divergence. Moreover, if D_{IFS} satisfies axiom IF-Div.5, then so does D^{φ}_{IFS}.

Proof Let D_{IFS} be an IF-divergence and φ an increasing function with $\varphi(0) = 0$. Condition IF-Diss.1 follows from $\varphi(0) = 0$ and conditions IF-Div.3 and IF-Div.4 follow from the monotonicity of φ, and IF-Div.2 and IF-Div.5 are trivially fulfilled by definition. ∎

5.1.3 Examples of IF-divergences and IF-dissimilarities

This subsection is devoted to the study of some of the most important examples of IF-divergences and dissimilarities. Specifically, we shall investigate whether the most prominent examples of dissimilarities that can be found in the literature are particular cases of IF-divergence. Furthermore, we shall also study if they satisfy other properties, such as axiom IF-Div.5, or if they are dissimilitudes.

Dissimilarities that also are IF-divergences

In this section we are going to present an overview of the dissimilarities that are also IF-divergences. From now on, Ω denotes a finite universe with n elements.

Hamming and normalized Hamming distance One of the most important comparison measures for IF-sets are the Hamming distance ([193]), defined by:

$$l_{IFS}(A, B) = \frac{1}{2} \sum_{\omega \in \Omega} (|\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)| + |\pi_A(\omega) - \pi_B(\omega)|),$$

and the normalized Hamming distance by:

$$l_{nIFS}(A,B) = \frac{1}{n} l_{IFS}(A,B), \quad \text{for every } A,B \in IFSs(\Omega).$$

These functions are known to be dissimilarities. Let us prove that they are also IF-divergences. In order to do this, we shall first of all prove that the Hamming distance is an IF-divergence; this, together with Proposition 5.27, we allow us to conclude that the normalized Hamming distance is also an IF-divergence, because it is an increasing transformation (by means of $\varphi(x) = \frac{x}{n}$) of the Hamming distance. In order to prove that the Hamming distance is an IF-divergence, we shall begin by showing that it satisfies axiom IF-Div.5. Let us note that

$$\pi_A(\omega) = 1 - \mu_A(\omega) - \nu_A(\omega) = 1 - \nu_{A^c}(\omega) - \mu_{A^c}(\omega) = \pi_{A^c}(\omega)$$

for every $\omega \in \Omega$ and $A \in IFSs(\Omega)$. Then:

$$\begin{aligned}
l_{IFS}(A^c,B^c) &= \sum_{\omega\in\Omega} \left(|\nu_A(\omega) - \nu_B(\omega)| + |\mu_A(\omega) - \mu_B(\omega)| + |\pi_{A^c}(\omega) - \pi_{B^c}(\omega)|\right) \\
&= \sum_{\omega\in\Omega} \left(|\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)| + |\pi_A(\omega) - \pi_B(\omega)|\right) = l_{IFS}(A,B).
\end{aligned}$$

By Proposition 5.23, axioms IF-Div.3 and IF-Div.4 are equivalent. Moreover, axioms IF-Diss.1 and IF-Diss.2 are satisfied since l_{IFS} is an IF-dissimilarity (see for instance [92]). Hence, in order to prove that l_{IFS} is an IF-divergence it suffices to check that it fulfills either IF-Div.3 or IF-Div.4. Let us show the latter. Let A, B and C be three IF-sets; using Lemma A.2 of Appendix A, we know that for every $\omega \in \Omega$, the following inequality holds:

$$\begin{aligned}
&|\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)| + |\pi_A(\omega) - \pi_B(\omega)| \geq \\
&\quad \left|\max(\mu_A(\omega),\mu_C(\omega)) - \max(\mu_B(\omega),\mu_C(\omega))\right| + \\
&\quad \left|\min(\nu_A(\omega),\nu_C(\omega)) - \min(\nu_B(\omega),\nu_C(\omega))\right| + \\
&\quad \big|\max(\mu_A(\omega),\mu_C(\omega)) + \min(\nu_A(\omega),\nu_C(\omega)) - \\
&\quad \max(\mu_B(\omega),\mu_C(\omega)) - \min(\nu_B(\omega),\nu_C(\omega))\big|.
\end{aligned}$$

Then:

$$\begin{aligned}
l_{IFS}(A,B) &= \sum_{\omega\in\Omega} |\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)| + |\pi_A(\omega) - \pi_B(\omega)| \\
&\geq \sum_{\omega\in\Omega} \big|\max(\mu_A(\omega),\mu_C(\omega)) - \max(\mu_B(\omega),\mu_C(\omega))\big| \\
&+ \big|\min(\nu_A(\omega),\nu_C(\omega)) - \min(\nu_B(\omega),\nu_C(\omega))\big| \\
&+ \big|\max(\mu_A(\omega),\mu_C(\omega)) - \min(\nu_A(\omega),\nu_C(\omega)) \\
&+ \max(\mu_B(\omega),\mu_C(\omega)) - \min(\nu_B(\omega),\nu_C(\omega))\big| = l_{IFS}(A \cup C, B \cup C).
\end{aligned}$$

Thus, $l_{IFS}(A,B) \geq l_{IFS}(A \cup C, B \cup C)$.

In other words, we have proven that l_{IFS} satisfies axiom IF-Div.4, and therefore it also satisfies IF-Div.3. Hence, l_{IFS} is an IF-divergence, and as a consequence so is l_{nIFS}.

Moreover, since they are IF-divergences, we deduce that they are also dissimilitudes. In summary, the Hamming and the normalized Hamming distances are examples of dissimilarities, IF-divergences, dissimilitudes and distances.

Hausdorff dissimilarity Another very important dissimilarity between IF-sets is based on the Hausdorff distance (see for example [85]). It is defined by:

$$d_H(A,B) = \sup_{\omega \in \Omega} \max(|\mu_A(\omega) - \mu_B(\omega)|, |\nu_A(\omega) - \nu_B(\omega)|).$$

As the Hamming distance, the Hausdorff dissimilarity satisfies axiom IF-Div.5, because

$$d_H(A^c, B^c) = \sup_{\omega \in \Omega} \max(|\nu_A(\omega) - \nu_B(\omega)|, |\mu_A(\omega) - \mu_B(\omega)|) = d_H(A, B).$$

Applying Prop 5.23, we deduce that axioms IF-Div.3 and IF-Div.4 are equivalent. Note that axioms IF-Diss.1 and IF-Diss.2 are satisfied by d_H since it is a IF-dissimilarity. Hence, in order to prove that d_H is an IF-divergence, it suffices to prove that either IF-Div.3 or IF-Div.4 hold.

Let us prove that axiom IF-Div.4 is satisfied by d_H. Consider three IF-sets A, B and C. Then, the IF-sets $A \cap C$ and $B \cap C$ are given by:

$$A \cap C = \{(\omega, \max(\mu_A(\omega), \mu_C(\omega)), \min(\nu_A(\omega), \nu_C(\omega))) \mid \omega \in \Omega\}.$$
$$B \cap C = \{(\omega, \max(\mu_B(\omega), \mu_C(\omega)), \min(\nu_B(\omega), \nu_C(\omega))) \mid \omega \in \Omega\}.$$

By the second part of Lemma A.1 of Appendix A, it holds that:

$$|\max(\mu_A(\omega), \mu_C(\omega)) - \max(\mu_B(\omega), \mu_C(\omega))| \leq |\mu_A(\omega) - \mu_B(\omega)|.$$
$$|\min(\nu_A(\omega), \nu_C(\omega)) - \min(\nu_B(\omega), \nu_C(\omega))| \leq |\nu_A(\omega) - \nu_B(\omega)|.$$

Then,

$$|\mu_{A\cap C}(\omega) - \mu_{B\cap C}(\omega)| \leq |\mu_A(\omega) - \mu_B(\omega)| \text{ and}$$
$$|\nu_{A\cap C}(\omega) - \nu_{B\cap C}(\omega)| \leq |\nu_A(\omega) - \nu_B(\omega)|.$$

From these inequalities it follows that:

$$\max(|\mu_{A\cap C}(\omega) - \mu_{B\cap C}(\omega)|, |\nu_{A\cap C}(\omega) - \nu_{B\cap C}(\omega)|)$$
$$\leq \max(|\mu_A(\omega) - \mu_B(\omega)|, |\nu_A(\omega) - \nu_B(\omega)|).$$

This inequality has been proved for every ω in Ω, and consequently:

$$\begin{aligned} d_H(A \cap C, B \cap C) &= \sup_{\omega \in \Omega} \max(|\mu_{A\cap C}(\omega) - \mu_{B\cap C}(\omega)|, |\nu_{A\cap C}(\omega) - \nu_{B\cap C}(\omega)|) \\ &\leq \sup_{\omega \in \Omega} \max(|\mu_A(\omega) - \mu_B(\omega)|, |\nu_A(\omega) - \nu_B(\omega)|) = d_H(A, B). \end{aligned}$$

Thus, the Hausdorff IF-dissimilarity is an IF-divergence, and consequently it is also a dissimilitude.

Note that it is also possible to define the normalized Hausdorff dissimilarity, denoted by d_{nH}, by:

$$d_{nH}(A,B) = \frac{1}{n} d_H(A,B), \text{ for every } A,B \in IFSs(\Omega).$$

It holds that $d_{nH}(A,B) = \varphi(d_H(A,B))$, where $\varphi(x) = \frac{1}{n}x$. As we already said, this function φ is increasing and $\varphi(0) = 0$. Therefore, using Proposition 5.27, we deduce that d_{nH} is also an IF-divergence that fulfills axiom IF-Div.5.

We conclude that d_H and d_{nH} are distances, IF-dissimilarities, IF-divergences and IF-dissimilitudes at the same time.

Hong & Kim dissimilarities Hong and Kim proposed two dissimilarity measures in [89]. They are defined by:

$$D_C(A,B) = \frac{1}{2n}\sum_{\omega\in\Omega}\left(|\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)|\right) \text{ and}$$

$$D_L(A,B) = \frac{1}{4n}\left(\sum_{\omega\in\Omega}|S_A(\omega) - S_B(\omega)| + \sum_{\omega\in\Omega}\left(|\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)|\right)\right),$$

where $S_A(\omega) = \mu_A(\omega) - \nu_A(\omega)$ and $S_B(\omega) = \mu_B(\omega) - \nu_B(\omega)$.

Recall that D_L can be equivalently expressed by:

$$D_L(A,B) = \frac{1}{4n}\sum_{\omega\in\Omega}\left(\left|(\mu_A(\omega) - \mu_B(\omega)) - (\nu_A(\omega) - \nu_B(\omega))\right| + |\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)|\right)$$

for every $A,B \in IFSs(\Omega)$.

In order to prove that D_C satisfies IF-Div.3, we shall use part b) of Lemma A.1:

$$|\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)| \geq$$
$$\left|\max(\mu_A(\omega),\mu_C(\omega)) - \max(\mu_B(\omega),\mu_C(\omega))\right| + \left|\min(\nu_A(\omega),\nu_C(\omega)) - \min(\nu_B(\omega),\nu_C(\omega))\right|.$$

Using this fact, IF-Div.3 trivially follows, and IF-Div.4 can be similarly proved.

Let us see that D_L is also an IF-divergence. For this, it suffices to take into account that, from Lemma A.3, for every $\omega \in \Omega$ it holds that:

$$\begin{aligned}
&|\mu_A(\omega) - \mu_B(\omega) - \nu_A(\omega) + \nu_B(\omega)| + |\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)| \\
&\quad\geq \big|\max(\mu_A(\omega),\mu_C(\omega)) - \max(\mu_B(\omega),\mu_C(\omega)) \\
&\quad - \min(\nu_A(\omega),\nu_C(\omega)) + \min(\nu_B(\omega),\nu_C(\omega))\big| \\
&\quad + \left|\max(\mu_A(\omega),\mu_C(\omega)) - \max(\mu_B(\omega),\mu_C(\omega))\right| \\
&\quad + \left|\min(\nu_A(\omega),\nu_C(\omega)) - \min(\nu_B(\omega),\nu_C(\omega))\right|.
\end{aligned}$$

By taking the sum on Ω on every part of the inequality, and multiplying each term by $\frac{1}{4n}$, we obtain that:

$$D_L(A,B) \geq D_L(A \cup C, B \cup C).$$

Thus, D_L satisfies axiom IF-Div.4, and therefore also IF-Div.3 since D_L satisfies the property IF-Div.5. We conclude that both D_C and D_L are IF-dissimilarities, IF-divergences and IF-dissimilitudes.

Li et al. dissimilarity Another dissimilarity measure for IF-sets was proposed by Li et al. ([113]):

$$D_O(A,B) = \left(\frac{1}{2n} \sum_{\omega \in \Omega} (\mu_A(\omega) - \mu_B(\omega))^2 + (\nu_A(\omega) - \nu_B(\omega))^2\right)^{\frac{1}{2}}.$$

This dissimilarity also satisfies IF-Div. 5, since $D_O(A^c, B^c) = D_O(A,B)$. Then, by Proposition 5.23, in order to prove that D_O is an IF-divergence it is enough to prove that it satisfies IF-Div.4. Let us consider A, B and C three IF-sets. By the second part of Lemma A.1 in Appendix A, we know that:

$$|\max(\mu_A(\omega), \mu_C(\omega)) - \max(\mu_B(\omega), \mu_C(\omega))| \leq |\mu_A(\omega) - \mu_B(\omega)| \text{ and}$$
$$|\min(\nu_A(\omega), \nu_C(\omega)) - \min(\nu_B(\omega), \nu_C(\omega))| \leq |\nu_A(\omega) - \nu_B(\omega)|,$$

or, equivalently,

$$|\mu_{A \cup C}(\omega) - \mu_{B \cup C}(\omega)| \leq |\mu_A(\omega) - \mu_B(\omega)| \text{ and}$$
$$|\nu_{A \cup C}(\omega) - \nu_{B \cup C}(\omega)| \leq |\nu_A(\omega) - \nu_B(\omega)|.$$

Then it holds that:

$$\begin{aligned} &|\mu_{A \cup C}(\omega) - \mu_{B \cup C}(\omega)|^2 + |\nu_{A \cup C}(\omega) - \nu_{B \cup C}(\omega)|^2 \\ &\qquad \leq |\mu_A(\omega) - \mu_B(\omega)|^2 + |\nu_A(\omega) - \nu_B(\omega)|^2, \end{aligned}$$

whence

$$\begin{aligned} D_O(A \cup C, B \cup C) &= \left(\frac{1}{2n} \sum_{\omega \in \Omega} |\mu_{A \cup C}(\omega) - \mu_{B \cup C}(\omega)|^2 + |\nu_{A \cup C}(\omega) - \nu_{B \cup C}(\omega)|^2\right)^{\frac{1}{2}} \\ &\leq \left(\frac{1}{2n} \sum_{\omega \in \Omega} |\mu_A(\omega) - \mu_B(\omega)|^2 + |\nu_A(\omega) - \nu_B(\omega)|^2\right)^{\frac{1}{2}} = D_O(A,B). \end{aligned}$$

Thus, D_O satisfies axiom IF-Div.4 and therefore it is an IF-Divergence, and in particular an IF-dissimilitude.

Mitchell dissimilarity Mitchell ([138]) proposed a dissimilarity defined by:

$$D_{HB}(A,B)=\frac{1}{2\sqrt[p]{n}}\left(\left(\sum_{\omega\in\Omega}|\mu_A(\omega)-\mu_B(\omega)|^p\right)^{\frac{1}{p}}+\left(\sum_{\omega\in\Omega}|\nu_A(\omega)-\nu_B(\omega)|^p\right)^{\frac{1}{p}}\right),$$

for some $p\geq 1$. This dissimilarity obviously satisfies IF-Div.5. Thus, in order to prove that D_{HB} is an IF-divergence it is enough to prove IF-Div.4, since IF-Diss.1 and IF-Diss.2 are satisfied for every dissimilarity. Consider A, B and C. Applying again the second part of Lemma A.1 from Appendix A we deduce that:

$$|\mu_{A\cap C}(\omega)-\mu_{B\cap C}(\omega)|\leq|\mu_A(\omega)-\mu_B(\omega)| \text{ and}$$
$$|\nu_{A\cap C}(\omega)-\nu_{B\cap C}(\omega)|\leq|\nu_A(\omega)-\nu_B(\omega)|.$$

Moreover, the inequalities holds if we raise every term to the power of p, whence

$$\begin{aligned}D_{HB}(A,B)&=\frac{1}{2\sqrt[p]{n}}\left(\left(\sum_{\omega\in\Omega}|\mu_{A\cap C}(\omega)-\mu_{B\cap C}(\omega)|^p\right)^{\frac{1}{p}}+\left(\sum_{\omega\in\Omega}|\nu_{A\cap C}(\omega)-\nu_{B\cap C}(\omega)|^p\right)^{\frac{1}{p}}\right)\\&\leq\frac{1}{2\sqrt[p]{n}}\left(\left(\sum_{\omega\in\Omega}|\mu_A(\omega)-\mu_B(\omega)|^p\right)^{\frac{1}{p}}+\left(\sum_{\omega\in\Omega}|\nu_A(\omega)-\nu_B(\omega)|^p\right)^{\frac{1}{p}}\right)\\&=D_{HB}(A,B).\end{aligned}$$

Thus, axiom IF-Div.4 holds, and therefore D_{HB} is an IF-divergence, and in particular a dissimilitude.

Liang & Shi dissimilarities Liang and Shi ([114]) defined the dissimilarities D_e^p and D_h^p, for some $p\geq 1$, by

$$D_e^p(A,B)=\frac{1}{2\sqrt[p]{n}}\left(\sum_{\omega\in\Omega}\big(|\mu_A(\omega)-\mu_B(\omega)|+|\nu_A(\omega)-\nu_B(\omega)|\big)^p\right)^{\frac{1}{p}},$$

$$D_h^p(A,B)=\frac{1}{\sqrt[p]{3n}}\left(\sum_{\omega\in\Omega}(\eta_1(\omega)+\eta_2(\omega)+\eta_3(\omega))^p\right)^{\frac{1}{p}},$$

where

$$\eta_1(\omega)=\tfrac{1}{2}(|\mu_A(\omega)-\mu_B(\omega)|+|\nu_A(\omega)-\nu_B(\omega)|).$$
$$\eta_2(\omega)=\tfrac{1}{2}|\mu_A(\omega)-\nu_A(\omega)-\mu_B(\omega)+\nu_B(\omega)|.$$
$$\eta_3(\omega)=\max(l_A(\omega),l_B(\omega))-\min(l_A(\omega),l_B(\omega)).$$
$$l_A(\omega)=\tfrac{1}{2}(1-\nu_A(\omega)-\mu_A(\omega)).$$
$$l_B(\omega)=\tfrac{1}{2}(1-\nu_B(\omega)-\mu_B(\omega)).$$

Note that D_h^p can be expressed in a equivalent way as

$$\begin{aligned}D_h^p(A,B)=\frac{1}{2\sqrt[p]{3n}}\Bigg(\sum_{\omega\in\Omega}\Big(&|\mu_A(\omega)-\mu_B(\omega)|+|\nu_A(\omega)-\nu_B(\omega)|\\&+\big|(\mu_A(\omega)-\mu_B(\omega))-(\nu_A(\omega)-\nu_B(\omega))\big|\\&+\big|(\mu_A(\omega)+\nu_A(\omega))-(\mu_B(\omega)+\nu_B(\omega))\big|\Big)^p\Bigg)^{\frac{1}{p}}.\end{aligned}$$

As in the previous examples, both D_e^p and D_h^p satisfy IF-Div. 5, and therefore it suffices to prove that both functions satisfy IF-Div.4 to prove that they are IF-divergences. Let us first focus on D_e^p, and let us consider A, B and C three IF-sets. Applying again the second part of Lemma A.1 in Appendix A we know that:

$$|\mu_{A\cup C}(\omega) - \mu_{B\cup C}(\omega)| \leq |\mu_A(\omega) - \mu_B(\omega)| \text{ and}$$
$$|\nu_{A\cup C}(\omega) - \nu_{B\cup C}(\omega)| \leq |\nu_A(\omega) - \nu_B(\omega)|.$$

If we sum both inequalities we obtain

$$|\mu_{A\cup C}(\omega) - \mu_{B\cup C}(\omega)| + |\nu_{A\cup C}(\omega) - \nu_{B\cup C}(\omega)| \leq |\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)|,$$

and since this inequality also holds when we raise every component to the power of p,

$$\begin{aligned} D_e^p(A\cup C, B\cup C) &= \frac{1}{\sqrt[p]{2n}} \left(\sum_{\omega\in\Omega} \left(|\mu_{A\cup C}(\omega) - \mu_{B\cup C}(\omega)| + |\nu_{A\cup C}(\omega) - \nu_{B\cup C}(\omega)|\right)^p\right)^{\frac{1}{p}} \\ &\leq \frac{1}{\sqrt[p]{2n}} \left(\sum_{\omega\in\Omega} \left(|\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)|\right)^p\right)^{\frac{1}{p}} = D_e^p(A, B). \end{aligned}$$

Thus, D_e^p satisfies IF-Div.4, and, taking into account that it satisfies IF-Div.5, also axiom IF-Div.3. Hence, it is a dissimilarity, and consequently, a dissimilitude.

Consider now D_h^p. Using Lemma A.4 in Appendix A, we know that, for every $\omega \in \Omega$,

$$\begin{aligned} &|\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)| + \\ &|\mu_A(\omega) - \mu_B(\omega) - \nu_A(\omega) + \nu_B(\omega)| + \\ &|\mu_A(\omega) + \nu_A(\omega) - \mu_B(\omega) - \nu_B(\omega)| \geq \\ &\qquad |\mu_{A\cup C}(\omega) - \mu_{B\cup C}(\omega)| + |\nu_{A\cup C}(\omega) - \nu_{B\cup C}(\omega)| + \\ &\qquad |\mu_{A\cup C}(\omega) - \mu_{B\cup C}(\omega) - \nu_{A\cup C}(\omega) + \nu_{B\cup C}(\omega)| + \\ &\qquad |\mu_{A\cup C}(\omega) + \nu_{A\cup C}(\omega) - \mu_{B\cup C}(\omega) - \nu_{B\cup C}(\omega)|. \end{aligned}$$

Making the summation over every ω in Ω in each part of the inequality and multiplying by $\frac{1}{\sqrt[p]{2\cdot 3n}}$, we obtain that $D_h^p(A, B) \geq D_h^p(A\cup C, B\cup C)$.

Thus, both D_e^p and D_h^p are IF-dissimilarities, IF-divergences and IF-dissimilitudes.

Hung & Yang dissimilarities Hung and Yang proposed some new dissimilarities in [92], two of which are based on the Hausdorff dissimilarity. As we shall see, it is easy to check that both are also IF-divergences. These dissimilarities are defined by:

$$\begin{aligned} D_{HY}^1(A, B) &= d_{nH}(A, B). \\ D_{HY}^2(A, B) &= 1 - \frac{e^{-d_{nH}(A,B)} - e^{-1}}{1 - e^{-1}}. \\ D_{HY}^3(A, B) &= 1 - \frac{1 - d_{nH}(A, B)}{1 + d_{nH}(A, B)}. \end{aligned}$$

We have already proven that the Hausdorff dissimilarity is an IF-divergence that satisfies the property IF-Div.5. Consider the functions φ_2 and φ_3 defined by:

$$\varphi_2(x) = 1 - \frac{e^{-x} - e^{-1}}{1 - e^{-1}} \text{ and } \varphi_3(x) = 1 - \frac{1 - x}{1 + x}.$$

These functions are increasing and satisfy $\varphi_2(0) = \varphi_3(0) = 0$. Applying Proposition 5.27 we conclude that

$$d_H^{\varphi_2}(A, B) = \varphi_2(d_{nH}(A, B)) = D_{HY}^2(A, B) \text{ and}$$
$$d_H^{\varphi_3}(A, B) = \varphi_3(d_{nH}(A, B)) = D_{HY}^3(A, B)$$

are IF-divergences that satisfy property IF-Div.5. Thus, they are also IF-dissimilitudes.

On the other hand, Hung and Yang also proposed the IF-dissimilarity given by

$$D_{pk2}(A, B) = \frac{1}{2}\left(\max_{\omega \in \Omega}|\mu_A(\omega) - \mu_B(\omega)| + \max_{\omega \in \Omega}|\nu_A(\omega) - \nu_B(\omega)|\right).$$

This measure satisfies IF-Div.5, whence, applying Proposition 5.23, it is enough to prove that, indeed, D_{pk2} satisfies IF-Div.4. If we consider A, B and C three IF-sets, we know from the second part of Lemma A.1 in Appendix A that:

$$|\mu_A(\omega) - \mu_B(\omega)| \ge |\max(\mu_A(\omega), \mu_C(\omega)) - \max(\mu_B(\omega), \mu_C(\omega))|.$$
$$|\nu_A(\omega) - \nu_B(\omega)| \ge |\min(\nu_A(\omega), \nu_C(\omega)) - \min(\nu_B(\omega), \nu_C(\omega))|.$$

Thus,

$$\max_{\omega \in \Omega}|\mu_A(\omega) - \mu_B(\omega)| \ge \max_{\omega \in \Omega}|\max(\mu_A(\omega), \mu_C(\omega)) - \max(\mu_B(\omega), \mu_C(\omega))|.$$
$$\max_{\omega \in \Omega}|\nu_A(\omega) - \nu_B(\omega)| \ge \max_{\omega \in \Omega}|\min(\nu_A(\omega), \nu_C(\omega)) - \min(\nu_B(\omega), \nu_C(\omega))|.$$

Then, $D_{pk2}(A, B) \ge D_{pk2}(A \cup C, B \cup C)$. We conclude that D_{pk2} is another example of IF-dissimilarity that is also an IF-divergence and IF-dissimilitude.

Dissimilarities that are not IF-divergences

Let us now provide some examples of dissimilarities, very frequently used in the literature, that are not IF-divergences. We shall also give some examples showing that these comparison measures are, in some cases, counterintuitive.

Euclidean and normalized Euclidean distance. Together with the Hamming and Hausdorff distances, one of the most important comparison measures is the Euclidean distance (see for example. This distance is used to define a dissimilarity between IF-sets and its normalization as follows ([85]):

$$q_{IFS}(A,B) = \left(\frac{1}{2} \sum_{\omega \in \Omega} (\mu_A(\omega) - \mu_B(\omega))^2 + (\nu_A(\omega) - \nu_B(\omega))^2 + (\pi_A(\omega) - \pi_B(\omega))^2 \right)^{\frac{1}{2}}.$$

$$q_{nIFS}(A,B) = \tfrac{1}{n} q_{IFS}(A,B).$$

These dissimilarities fulfill axiom IF-Div.5, since $\pi_A(\omega) = \pi_{A^c}(\omega)$ and $\pi_B(\omega) = \pi_{B^c}(\omega)$ for every $A, B \in IFSs(\Omega)$. However, they are not IF-divergences, since they do not satisfy axioms IF-Div.3 nor IF-Div.4. To see a counterexample, consider $\Omega = \{\omega\}$ and the following IF-sets:

$$A = \{(\omega, 0.12, 0.68)\} \quad B = \{(\omega, 0.29, 0.59)\} \quad C = \{(\omega, 0.11, 0.36)\}$$

The IF-sets $A \cup C$ and $B \cup C$ are given by:

$$A \cup C = \{(\omega, 0.12, 0.36)\} \text{ and } B \cup C = \{(\omega, 0.29, 0.36)\}$$

It holds that $q_{IFS}(A \cup C, B \cup C) > q_{IFS}(A,B)$:

$$q_{IFS}(A \cup C, B \cup C) = \left(\tfrac{1}{2}(0.17^2 + 0 + 0.17^2)\right)^{0.5} = 0.17.$$
$$q_{IFS}(A,B) = \left(\tfrac{1}{2}(0.17^2 + 0.09^2 + 0.08^2)\right)^{0.5} = 0.1473.$$

Moreover, since q_{IFS} does not satisfy IF-Div.4, axiom IF-Div.3 cannot hold either because they are equivalent under IF-Div.5. Therefore, q_{IFS} is neither an IF-divergence nor a dissimilitude. The same example shows that q_{nIFS} is not an IF-divergence, since for $n = 1$ we have that $q_{IFS} = q_{nIFS}$.

Liang & Shi dissimilarity We have seen previously some IF-dissimilarities proposed by Liang and Shi that are also IF-divergences. They also proposed another IF-dissimilarity measure, that is defined by:

$$D_s^p(A,B) = \frac{1}{\sqrt[p]{n}} \left(\sum_{\omega \in \Omega} (\phi_{s1}(\omega) + \phi_{s2}(\omega))^p \right)^{\frac{1}{p}},$$

where $p \geq 1$ and

$$\phi_{s1}(\omega) = \tfrac{1}{2}|m_{A1}(\omega) - m_{B1}(\omega)|.$$
$$\phi_{s2}(\omega) = \tfrac{1}{2}|m_{A2}(\omega) - m_{B2}(\omega)|.$$
$$m_{A1}(\omega) = \tfrac{1}{2}(\mu_A(\omega) + m_A(\omega)).$$
$$m_{A2}(\omega) = \tfrac{1}{2}(m_A(\omega) + 1 - \nu_A(\omega)).$$
$$m_{B1}(\omega) = \tfrac{1}{2}(\mu_B(\omega) + m_B(\omega)).$$
$$m_{B2}(\omega) = \tfrac{1}{2}(m_B(\omega) + 1 - \nu_B(\omega)).$$
$$m_A(\omega) = \tfrac{1}{2}(\mu_A(\omega) + 1 - \nu_A(\omega)).$$
$$m_B(\omega) = \tfrac{1}{2}(\mu_B(\omega) + 1 - \nu_B(\omega)).$$

Note that D^p_s can also be expressed by:

$$D^p_s(A,B) = \frac{1}{\sqrt[p]{n}}\left(\sum_{\omega\in\Omega}\frac{1}{8}\Big(\big|3(\mu_A(\omega)-\mu_B(\omega)) - (\nu_A(\omega)-\nu_B(\omega))\big|^p + \big|(\mu_A(\omega)-\mu_B(\omega)) - 3(\nu_A(\omega)-\nu_B(\omega))\big|^p\Big)\right)^{\frac{1}{p}}.$$

Thus, this dissimilarity satisfies axiom IF-Div.5. However, neither IF-Div.3 nor IF-Div.4 are satisfied. To see this, consider $\Omega=\{\omega\}$ and the IF-sets

$$A=\{(\omega, 0.25, 0.25)\} \text{ and } B=\{(\omega, 0.6, 0.35)\}.$$

For these IF-sets it holds that $D^p_s(A,B)=0.125$. Furthermore, if we consider the IF-set C defined by:

$$C=\{(\omega, 0.2, 0.2)\}$$

it holds that

$$A\cup C=\{(\omega, 0.25, 0.2)\} \text{ and } B\cup C=\{(\omega, 0.6, 0.2)\}$$

whence,

$$D^p_s(A\cup C, B\cup C)=0.175>0.125=D(A,B).$$

Consequently, D^p_s is neither an IF-divergence, nor an IF-dissimilitude.

Chen dissimilarity Chen ([36, 37]) defined an IF-dissimilarity measure by:

$$D_C(A,B)=\frac{1}{2n}\sum_{\omega\in\Omega}|S_A(\omega)-S_B(\omega)|,$$

where $S_A(\omega)=\mu_A(\omega)-\nu_A(\omega)$ and $S_B(\omega)=\mu_B(\omega)-\nu_B(\omega)$.

This dissimilarity also satisfies axiom IF-Div.5, because:

$$\begin{aligned}D_C(A^c,B^c) &= \frac{1}{2n}\sum_{\omega\in\Omega}|S_{A^c}(\omega)-S_{B^c}(\omega)| \\ &= \frac{1}{2n}\sum_{\omega\in\Omega}|\mu_A(\omega)-\mu_B(\omega)-\nu_A(\omega)+\nu_B(\omega)| \\ &= \frac{1}{2n}\sum_{\omega\in\Omega}|S_A(\omega)-S_B(\omega)| = D(A,B).\end{aligned}$$

By Proposition 5.23 axioms IF-Div.3 and IF-Div.4 are equivalent. Let us see an example where axiom IF-Div.4 is violated. Consider $\Omega=\{\omega\}$ and the IF-sets:

$$A=\{(\omega, 0.25, 0.75)\} \text{ and } B=\{(\omega, 0, 0.5)\}$$

It holds that $D_C(A,B)=0$. If we consider $C=\{(\omega, 0.2, 0.6)\}$, it holds that:

$$A\cup C=\{\omega, 0.25, 0.6\} \text{ and } B\cup C=\{\omega, 0.2, 0.5\}$$

whence

$$D_C(A \cup C, B \cup C) = 0.025 > 0 = D_C(A, B).$$

Thus, D_C is neither an IF-divergence nor a dissimilitude.

In [89], Hong provided an example that showed that this IF-dissimilarity is a counterintuitive measure of comparison of fuzzy sets. The main reason is that:

$$\mu_A(\omega) - \nu_A(\omega) = \mu_B(\omega) - \nu_B(\omega) \quad \forall \omega \in \Omega \Rightarrow D_C(A, B) = 0.$$

In fact, if we consider the IF-sets A and B defined by:

$$A = \{(\omega, 0, 0) \mid \omega \in \Omega\} \text{ and } B = \{(\omega, 0.5, 0.5) \mid \omega \in \Omega\};$$

we obtain $D_C(A, B) = 0$. However, these IF-sets do not seem to be very similar.

Dengfenf & Chuntian dissimilarity Dengfenf and Chuntian ([111]) proposed the following IF-dissimilarity:

$$D_{DC}(A, B) = \frac{1}{\sqrt[p]{n}} \sqrt[p]{\sum_{\omega \in \Omega} \left| \frac{1}{2}(\mu_A(\omega) - \mu_B(\omega) - \nu_A(\omega) + \nu_B(\omega)) \right|^p},$$

for some $p \geq 1$. Again, it obviously holds that $D(A^c, B^c) = D(A, B)$, that is, D_{DC} satisfies IF-Div.5, and therefore, by Proposition 5.23, axioms IF-Div.3 and IF-Div.4 are equivalent. Furthermore, when $p = 1$, D_{DC} becomes Chen dissimilarity multiplied by a constant. Thus, in order to obtain a counterexample, it suffices to consider the same than in the previous paragraph.

Hung & Yang dissimilarities Previously we have seen some examples of IF-dissimilarities proposed by Hung and Yang that are also IF-divergences. Here we give some examples of IF-dissimilarities proposed by them which are not IF-divergences. They are given by:

$$D_{\omega 1}(A, B) = 1 - \frac{1}{n} \sum_{\omega \in \Omega} \frac{\min(\mu_A(\omega), \mu_B(\omega)) + \min(\nu_A(\omega), \nu_B(\omega))}{\max(\mu_A(\omega), \mu_B(\omega)) + \max(\nu_A(\omega), \nu_B(\omega))}.$$

$$D_{pk1}(A, B) = 1 - \frac{\sum_{\omega \in \Omega} \min(\mu_A(\omega), \mu_B(\omega)) + \min(\nu_A(\omega), \nu_B(\omega))}{\sum_{\omega \in \Omega} \max(\mu_A(\omega), \mu_B(\omega)) + \max(\nu_A(\omega), \nu_B(\omega))}.$$

$$D_{pk3}(A, B) = \frac{\sum_{\omega \in \Omega} |\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)|}{\sum_{\omega \in \Omega} |\mu_A(\omega) + \mu_B(\omega)| + |\nu_A(\omega) + \nu_B(\omega)|}.$$

These dissimilarities satisfy axiom IF-Div.5, and therefore, using Proposition 5.23, both axioms IF-Div.3 and IF-Div.4 become equivalent. However, none of them satisfies these axioms. Let us give a counterexample for $D_{\omega 1}$: consider an universe $\Omega = \{\omega\}$ and the IF-sets:

$$A = \{(\omega, 0.75, 0.19)\} \text{ and } B = \{(\omega, 0.48, 0.23)\}$$

For these IF-sets, $D_{\omega 1}(A, B) = 0.32$. If we now consider the IF-set $C = \{(\omega, 0.25, 0.06)\}$ then $A \cup C$ and $B \cup C$ are given by:

$$A \cup C = \{(\omega, 0.75, 0.06)\} \text{ and } B \cup C = \{(\omega, 0.48, 0.06)\}$$

Hence,

$$D_{\omega 1}(A \cup C, B \cup C) \geq 0.333 > 0.32 = D(A, B).$$

The same example shows that D_{pk1} does not satisfy IF-Div.4, since for $n = 1$ D_{pk1} and $D_{\omega 1}$ are the same function.

Let us prove now that $D_{\omega 3}$ does not satisfy IF-Div.4 neither. For this, take $\Omega = \{\omega\}$ define the following IF-sets:

$$A = \{(\omega, 0.24, 0.28)\},\ B = \{(\omega, 0.66, 0.29)\},\ C = \{(\omega, 0.02, 0.15)\}$$

Then, it holds that:

$$D_{pk3}(A, B) = 0.29 < 0.35 = D_{pk3}(A \cup C, B \cup C).$$

Thus, none of these IF-dissimilarity measures are IF-divergences or IF-dissimilitudes.

In Table 5.1 we have summarized the results we have presented in this section. There, we can see which axioms satisfy every one of the examples of IF-dissimilarities we have studied. We can remark that all these examples satisfy the property IF-Div.5, and then IF-Div.3 and IF-Div.4 are equivalent. Recall that all the measures we have studied satisfy property IF-Div.5, and then IF-divergences and IF-dissimilitudes become equivalent.

5.1.4 Local IF-divergences

In this section we are going to study a special type of IF-divergences called the local IF-divergences. They are an important family of IF-divergences because of the interesting properties they satisfy.

Let us consider an universe $\Omega = \{\omega_1, \ldots, \omega_n\}$ and an IF-divergence D_{IFS} defined on $IFSs(\Omega) \times IFSs(\Omega)$. From IF-Div.4, we know that $D(A \cup C, B \cup C) \leq D(A, B)$ for every $C \in IFSs(\Omega)$. In particular, given $C = \{\omega_i\}$, we can express it equivalently by

$$C = \{(\omega_i, 1, 0), (\omega_j, 0, 1) \mid j \neq i\}.$$

Name	Notation	IF-Diss.1&2	IF-Div.3&4	IF-Div.5	IF-diss	IF-div
Hamming	l_{IFS}	OK	OK	OK	Yes	Yes
Normalized Hamming	l_{nIFS}	OK	OK	OK	Yes	Yes
Hausdorff	d_H	OK	OK	OK	Yes	Yes
Normalized Hausdorff	d_{nH}	OK	OK	OK	Yes	Yes
Normalized Eucliden	q_{IFS}	OK	FAIL	OK	Yes	No
Hong and Kim (I)	D_C	OK	OK	OK	Yes	Yes
Hong and Kim (I I)	D_L	OK	OK	OK	Yes	Yes
Li et al.	D_O	OK	OK	OK	Yes	Yes
Mitchell	D_{HB}	OK	OK	OK	Yes	Yes
Liang and Shi (I)	D_e^p	OK	OK	OK	Yes	Yes
Liang and Shi (I I)	D_h^p	OK	OK	OK	Yes	Yes
Liang and Shi (I I I)	D_s^p	OK	FAIL	OK	Yes	No
Chen	D_C	OK	FAIL	OK	Yes	No
Dengfeng and Chuntian	D_{DC}	OK	FAIL	OK	Yes	No
Hung and Yang (I)	D_{HY}^1	OK	OK	OK	Yes	Yes
Hung and Yang (I I)	D_{HY}^2	OK	OK	OK	Yes	Yes
Hung and Yang (I I I)	D_{HY}^3	OK	OK	OK	Yes	Yes
Hung and Yang (IV)	$D_{\omega 1}$	OK	FAIL	OK	Yes	No
Hung and Yang (V)	D_{pk1}	OK	FAIL	OK	Yes	No
Hung and Yang (VI)	D_{pk2}	OK	OK	OK	Yes	Yes
Hung and Yang (VI I I)	D_{pk3}	OK	FAIL	OK	Yes	No

Table 5.1: Behaviour of well-know n dissimilarities and IF-divergences.

Then, the IF-sets $A \cup \{\omega_i\}$ and $B \cup \{\omega_i\}$ are given by:

$$A \cup \{\omega_i\} = \{(\omega_i, 1, 0), (\omega_j, \mu_A(\omega_j), \nu_A(\omega_j)) \mid j \neq i\}.$$
$$B \cup \{\omega_i\} = \{(\omega_i, 1, 0), (\omega_j, \mu_B(\omega_j), \nu_B(\omega_j)) \mid j \neq i\}.$$

Applying axiom IF-Div.4 to these IF-sets, we obtain the following inequality:

$$D_{\mathrm{IFS}}(A \cup \{\omega_i\}, B \cup \{\omega_i\}) \leq D_{\mathrm{IFS}}(A, B).$$

Hence, the only difference between $D_{\mathrm{IFS}}(A \cup C, B \cup C)$ and $D_{\mathrm{IFS}}(A, B)$ is on the i-th element. However, such a function may not exist. When it does, the IF-divergence will be called local.

Definition 5.28 *Let D_{IFS} be an IF-divergence. It is called local (or it is said to satisfy the local property) when for every $A, B \in IFSs(\Omega)$ and every $\omega \in \Omega$ it holds that:*

$$D_{\mathrm{IFS}}(A, B) - D_{\mathrm{IFS}}(A \cup \{\omega\}, B \cup \{\omega\}) = h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega)). \quad (5.3)$$

In order to characterize local IF-divergences we are going to see the next Theorem.

Theorem 5.29 *A map $D_{\mathrm{IFS}} : IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ on a finite universe $\Omega = \{\omega_1, \ldots, \omega_n\}$ is a local IF-divergence if and only if there is a function $h_{IFS} : T^2 \to \mathbb{R}$ such that for every $A, B \in IFSs(\Omega)$:*

$$D_{\mathrm{IFS}}(A, B) = \sum_{i=1}^{n} h_{IFS}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)), \quad (5.4)$$

where T denotes the set $T = \{(t, z) \in [0, 1]^2 \mid t + z \leq 1\}$ and h_{IFS} fulfils the following properties:

- *IF-loc.1* $h_{IFS}(x, y, x, y) = 0$ *for every* $(x, y) \in T$.
- *IF-loc.2* $h_{IFS}(x_1, x_2, y_1, y_2) = h_{IFS}(y_1, y_2, x_1, x_2)$ *for every* $(x_1, x_2), (y_1, y_2) \in T$.
- *IF-loc.3* *If* $(x_1, x_2), (y_1, y_2) \in T$, $z \in [0, 1]$ *and* $x_1 \leq z \leq y_1$, *it holds that:* $h_{IFS}(x_1, x_2, y_1, y_2) \geq h_{IFS}(x_1, x_2, z, y_2)$. *Moreover, if* $(x_2, z), (y_2, z) \in T$ *it holds that* $h_{IFS}(x_1, x_2, y_1, y_2) \geq h_{IFS}(z, x_2, y_1, y_2)$.
- *IF-loc.4* *If* $(x_1, x_2), (y_1, y_2) \in T$, $z \in [0, 1]$ *and* $x_2 \leq z \leq y_2$, *it holds that:* $h_{IFS}(x_1, x_2, y_1, y_2) \geq h_{IFS}(x_1, x_2, y_1, z)$. *Moreover, if* $(x_1, z), (y_1, z) \in T$ *it holds that:* $h_{IFS}(x_1, x_2, y_1, y_2) \geq h_{IFS}(x_1, z, y_1, y_2)$.
- *IF-loc.5* *If* $(x_1, x_2), (y_1, y_2) \in T$ *and* $z \in [0, 1]$ *then:* $h_{IFS}(z, x_2, z, y_2) \leq h_{IFS}(x_1, x_2, y_1, y_2)$ *if* $(x_2, z), (y_2, z) \in T$ *and* $h_{IFS}(x_1, z, y_1, z) \leq h_{IFS}(x_1, x_2, y_1, y_2)$ *if* $(x_1, z), (y_1, z) \in T$.

Proof Assume first of all that D_{IFS} is a local IF-divergence and let us prove that $D_{IFS}(A, B)$ can be expressed as in Equation (5.4) for every $A, B \in IFSs(\Omega)$, where h_{IFS} satisfies the properties IF-loc.1 to IF-loc.6. In order to prove that, we will apply recursively Equation (5.3):

$$\begin{aligned} D_{IFS}(A, B) &= D_{IFS}(A \cup \{\omega_1\}, B \cup \{\omega_1\}) \\ &\quad + h_{IFS}(\mu_A(\omega_1), \nu_A(\omega_1), \mu_B(\omega_1), \nu_B(\omega_1)) \\ &= D_{IFS}(A \cup \{\omega_1\} \cup \{\omega_2\}, B \cup \{\omega_1\} \cup \{\omega_2\}) \\ &\quad + \sum_{i=1}^{2} h_{IFS}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)) \\ &= \ldots \\ &= D_{IFS}(\Omega, \Omega) + \sum_{i=1}^{n} h_{IFS}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)). \end{aligned}$$

Moreover, from axiom IF-Diss.1 we know that $D_{IFS}(\Omega, \Omega) = 0$, and therefore D_{IFS} can be expressed by:

$$D_{IFS}(A, B) = \sum_{i=1}^{n} h_{IFS}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)).$$

This shows that D_{IFS} can be expressed as in Equation (5.4).

Let us prove next that h_{IFS} fulfills properties IF-loc.1 to IF-loc.5:

~~IF-loc.1~~: Take $x, y \in T$, and let us prove that $h_{IFS}(x, y, x, y) = 0$. Define the IF-set A by $\mu_A(\omega_i) = x$ and $\nu_A(\omega_i) = y$, for every $i = 1, \ldots, n$. Note that A is in fact an IF-set since $\mu_A(\omega_i) + \nu_A(\omega_i) = x + y \leq 1$ for every $i = 1, \ldots, n$. Applying IF-diss.1, $D_{IFS}(A, A) = 0$, and therefore, since $D_{IFS}(A, A)$ can be expressed as in Equation (5.4), it holds that:

$$\begin{aligned} 0 = D_{IFS}(A, A) &= \sum_{i=1}^{n} h_{IFS}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_A(\omega_i), \nu_A(\omega_i)) \\ &= \sum_{i=1}^{n} h_{IFS}(x, y, x, y) = n \cdot h_{IFS}(x, y, x, y). \end{aligned}$$

Then, it must hold that $h_{IFS}(x, y, x, y) = 0$.

~~IF-loc.2~~: Let $(x_1, x_2), (y_1, y_2)$ be two elements in T. Consider the IF-sets A and B defined by: $\mu_A(\omega_i) = x_1$, $\nu_A(\omega_i) = x_2$, $\mu_B(\omega_i) = y_1$ and $\nu_B(\omega_i) = y_2$. Using axiom

IF-diss.2 and Equation(5.4) we obtain the following:

$$
\begin{aligned}
n\, h_{\mathrm{IFS}}(x_1,x_2,y_1,y_2) &= \sum_{i=1}^{n} h_{\mathrm{IFS}}(\mu_A(\omega_i),\nu_A(\omega_i),\mu_B(\omega_i),\nu_B(\omega_i)) \\
&= D_{\mathrm{IFS}}(A,B) = D_{\mathrm{IFS}}(B,A) \\
&= \sum_{i=1}^{n} h_{\mathrm{IFS}}(\mu_B(\omega_i),\nu_B(\omega_i),\mu_A(\omega_i),\nu_A(\omega_i)) \\
&= n h_{\mathrm{IFS}}(y_1,y_2,x_1,x_2).
\end{aligned}
$$

Thus, $h_{\mathrm{IFS}}(x_1,x_2,y_1,y_2) = h_{\mathrm{IFS}}(y_1,y_2,x_1,x_2)$.

~~IF-loc.3~~: Consider $(x_1,x_2), (y_1,y_2) \in T$ and $z \in [0,1]$ such that $x_1 \le z \le y_1$, and let us define the IF-sets A and B by: $\mu_A(\omega_i) = x_1$, $\nu_A(\omega_i) = x_2$, $\mu_B(\omega_i) = y_1$ and $\nu_B(\omega_i) = y_2$, for every $i = 1,\ldots,n$. We have to consider two cases:

- On one hand we are going to prove that

$$h_{\mathrm{IFS}}(x_1,x_2,y_1,y_2) \ge h_{\mathrm{IFS}}(x_1,x_2,z,y_2).$$

  To see this, consider the IF-set C defined by $\mu_C(\omega_i) = z$ and $\nu_C(\omega_i) = 0$ for $i = 1,\ldots,n$. Then the IF-sets $A \cap C$ and $B \cap C$ are given by:

$$
\begin{aligned}
A \cap C &= A. \\
B \cap C &= \{(\omega_i,\mu_C(\omega_i),\nu_B(\omega_i)) \mid i = 1,\ldots,n\}.
\end{aligned}
$$

  By axiom IF-Div.3, we see that $D_{\mathrm{IFS}}(A,B) \ge D_{\mathrm{IFS}}(A \cap C, B \cap C) = D_{\mathrm{IFS}}(A, B \cap C)$, and then Equation (5.4) implies that:

$$
\begin{aligned}
n\, h_{\mathrm{IFS}}(x_1,x_2,y_1,y_2) = D_{\mathrm{IFS}}(A,B) &\ge D_{\mathrm{IFS}}(A \cap C, B \cap C) \\
&= n\, h_{\mathrm{IFS}}(x_1,x_2,z,y_2).
\end{aligned}
$$

  Hence, $h_{\mathrm{IFS}}(x_1,x_2,y_1,y_2) \ge h_{\mathrm{IFS}}(x_1,x_2,z,y_2)$.

- Let us prove now that, when $(x_2,z), (y_2,z) \in T$, it holds that $h_{\mathrm{IFS}}(x_1,x_2,y_1,y_2) \ge h_{\mathrm{IFS}}(z,x_2,y_1,y_2)$. Consider the IF-set C defined by $\mu_C(\omega_i) = z$ and $\nu_C(\omega_i) = \max(x_2,y_2)$, for $i = 1,\ldots,n$. Note that C is an IF-set because $\mu_C(\omega_i) + \nu_C(\omega_i) = \max(x_2+z, y_2+z) \le 1$, for $i = 1,\ldots,n$. Using axiom IF-Div.4, we deduce that $D_{\mathrm{IFS}}(A,B) \ge D_{\mathrm{IFS}}(A \cup C, B \cup C)$. Moreover, the IF-sets $A \cup C$ and $B \cup C$ are given by:

$$
\begin{aligned}
A \cup C &= \{(\omega_i,\mu_C(\omega_i),\nu_A(\omega_i)) \mid i = 1,\ldots,n\}. \\
B \cup C &= B.
\end{aligned}
$$

  Then, $D_{\mathrm{IFS}}(A,B) \ge D_{\mathrm{IFS}}(A \cup C, B)$. This, together with Equation (5.4), implies that:

$$n\, h_{\mathrm{IFS}}(x_1,x_2,y_1,y_2) = D_{\mathrm{IFS}}(A,B) \ge D_{\mathrm{IFS}}(A \cup C, B \cup C) = n\, h_{\mathrm{IFS}}(z,x_2,y_1,y_2).$$

  Hence, $h_{\mathrm{IFS}}(x_1,x_2,y_1,y_2) \ge h_{\mathrm{IFS}}(z,x_2,y_1,y_2)$.

~~IF-loc.4~~: The proof is similar to that of IF-loc.3. Consider (x_1, x_2) and (y_1, y_2) in T, and let z be a point in $[0, 1]$ such that $x_2 \leq z \leq y_2$. Define the IF-sets A and B by:

$$A = \{(\omega, x_1, x_2) \mid \omega \in \Omega\} \text{ and } B = \{(\omega, y_1, y_2) \mid \omega \in \Omega\}.$$

If we consider the IF-set C given by:

$$C = \{(\omega, 0, z) \mid \omega \in \Omega\},$$

then, the IF-sets $A \cup C$ and $B \cup C$ are given by:

$$A \cup C = A \text{ and } B \cup C = \{(\omega, y_1, z)\}.$$

Applying axiom IF-Div.4 we deduce that

$$D_{\text{IFS}}(A, B) \geq D_{\text{IFS}}(A \cup C, B \cup C) = D_{\text{IFS}}(A, B \cup C),$$

and using now Equation (5.4), we obtain:

$$n\, h_{\text{IFS}}(x_1, x_2, y_1, y_2) = D_{\text{IFS}}(A, B) \geq D_{\text{IFS}}(A \cup C, B \cup C) = n\, h_{\text{IFS}}(x_1, x_2, y_1, z).$$

Moreover, if $(x_1, z), (y_1, z) \in T$, we consider the set

$$C = \{(\omega, \max(x_1, y_1), z) \mid \omega \in \Omega\}.$$

Since $(x_1, z), (y_1, z) \in T$, C is an IF-set. Moreover, $A \cap C$ and $B \cap C$ are given by:

$$A \cap C = \{(\omega, x_1, z) \mid \omega \in \Omega\} \text{ and } B \cap C = B.$$

Using axiom IF-Div.3, we deduce that

$$D_{\text{IFS}}(A, B) \geq D_{\text{IFS}}(A \cap C, B \cap C) = D_{\text{IFS}}(A \cap C, B),$$

and applying Equation (5.4),

$$n\, h_{\text{IFS}}(x_1, x_2, y_1, y_2) = D_{\text{IFS}}(A, B) \geq D_{\text{IFS}}(A \cap C, B \cap C) = n\, h_{\text{IFS}}(x_1, z, y_1, y_2).$$

Hence, $h_{\text{IFS}}(x_1, x_2, y_1, y_2) \geq h_{\text{IFS}}(x_1, z, y_1, y_2)$.

~~IF-loc.5~~: Let us consider $(x_1, x_2), (y_1, y_2) \in T$ and $z \in [0, 1]$.

If we assume that $(x_2, z), (y_2, z) \in T$, then $\max(x_2, y_2) + z \leq 1$; we consider the IF-sets A, B, C and D given by:

$$A = \{(\omega, x_1, x_2) \mid \omega \in \Omega\}, \quad B = \{(\omega, y_1, y_2) \mid \omega \in \Omega\}.$$
$$C = \{(\omega, z, x_2) \mid \omega \in \Omega\}, \quad D = \{(\omega, z, y_2) \mid \omega \in \Omega\}.$$

From Proposition 5.25, we know that $D_{\text{IFS}}(A, B) \geq D_{\text{IFS}}(C, D)$, and applying Equation (5.4) we deduce that

$$n\, h_{\text{IFS}}(x_1, x_2, y_1, y_2) = D_{\text{IFS}}(A, B) \geq D_{\text{IFS}}(C, D) = n\, h_{\text{IFS}}(z, x_2, z, y_2).$$

Thus, $h_{IFS}(x_1, x_2, y_1, y_2) \geq h_{IFS}(z, x_2, z, y_2)$.

If we assume now that $(x_1, z), (y_1, z)$ T, it holds that $\max(x_1, y_1) + z \leq 1$; we consider the IF-sets:

$$A = \{(\omega, x_1, x_2) \mid \omega \quad \Omega\}, \qquad B = \{(\omega, y_1, y_2) \mid \omega \quad \Omega\}.$$
$$C = \{(\omega, x_1, z) \mid \omega \quad \Omega\}, \qquad D = \{(\omega, y_1, z) \mid \omega \quad \Omega\}.$$

Applying Corollary 5.26, $D_{IFS}(A, B) \geq D_{IFS}(C, D)$. Using Equation (5.4), we obtain:

$$n \; h_{IFS}(x_1, x_2, y_1, y_2) = D_{IFS}(A, B) \geq D_{IFS}(C, D) = n \; h_{IFS}(x_1, z, y_1, z).$$

Thus, $h_{IFS}(x_1, x_2, y_1, y_2) \geq h_{IFS}(x_1, z, y_1, z)$.

Summarizing, if D_{IFS} is a local IF-divergence, then $D_{IFS}(A, B)$ can be expressed as in Equation (5.4) where the function h_{IFS} satisfies IF-loc.1 to IF-loc.5.

Let us prove the converse: that if a function D_{IFS} is defined by Equation (5.4), where h_{IFS} fulfills properties IF-loc.1 to IF-loc.5, then D_{IFS} is a local IF-divergence.

First of all, let us prove that D_{IFS} is an IF-divergence, i.e., that it satisfies axioms IF-Diss.1, IF-Diss.2, IF-Div.3 and IF-Div.4.

~~IF-Diss.1:~~ Let A be an IF-set. Then, $D_{IFS}(A, A) = 0$ because

$$D_{IFS}(A, A) = \sum_{i=1}^{n} h_{IFS}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_A(\omega_i), \nu_A(\omega_i)) = 0,$$

since IF-loc.1 implies that $h_{IFS}(x, y, x, y) = 0$ for every (x, y) T, and in particular $(\mu_A(\omega_i), \nu_A(\omega_i))$ T.

~~IF-Diss.2:~~ Let A, B be IF-sets, and let us prove that $D_{IFS}(A, B) = D_{IFS}(B, A)$. By IF-loc.2, $h_{IFS}(x_1, x_2, y_1, y_2) = h_{IFS}(y_1, y_2, x_1, x_2)$ for every $(x_1, x_2), (y_1, y_2)$ T, as $(\mu_A(\omega_i), \nu_A(\omega_i)), (\mu_B(\omega_i), \nu_B(\omega_i))$ T, whence

$$D_{IFS}(A, B) = D_{IFS}(B, A).$$

~~IF-Div.3 & IF-Div.4:~~ Consider three IF-sets A, B and C, and let us show that $D_{IFS}(A, B) \geq \max(D_{IFS}(A \quad C, B \quad C), D_{IFS}(A \cap C, B \cap C))$. Consider the following partition of Ω:

$$P_1 = \{\omega \quad \Omega \mid \max(\mu_A(\omega), \mu_B(\omega)) \leq \mu_C(\omega)\}.$$
$$P_2 = \{\omega \quad \Omega \mid \mu_A(\omega) \leq \mu_C(\omega) < \mu_B(\omega)\}.$$
$$P_3 = \{\omega \quad \Omega \mid \mu_B(\omega) \leq \mu_C(\omega) < \mu_A(\omega)\}.$$
$$P_4 = \{\omega \quad \Omega \mid \mu_C(\omega) < \min(\mu_A(\omega), \mu_B(\omega))\}.$$
$$Q_1 = \{\omega \quad \Omega \mid \max(\nu_A(\omega), \nu_B(\omega)) \leq \nu_C(\omega)\}.$$
$$Q_2 = \{\omega \quad \Omega \mid \nu_A(\omega) \leq \nu_C(\omega) < \nu_B(\omega)\}.$$
$$Q_3 = \{\omega \quad \Omega \mid \nu_B(\omega) \leq \nu_C(\omega) < \nu_A(\omega)\}.$$
$$Q_4 = \{\omega \quad \Omega \mid \nu_C(\omega) < \min(\nu_A(\omega), \nu_B(\omega))\}.$$

Thus, $\Omega = \bigcup_{i=1}^{4}\bigcup_{j=1}^{4}(P_i \cap Q_j)$. We are going to prove that, for every $i,j \in \{1,\ldots,4\}$, if $\omega \in P_i \cap Q_j$ then both:

$$h_{\text{IFS}}(\mu_{A\cup C}(\omega), \nu_{A\cup C}(\omega), \mu_{B\cup C}(\omega), \nu_{B\cup C}) \text{ and}$$
$$h_{\text{IFS}}(\mu_{A\cap C}(\omega), \nu_{A\cap C}(\omega), \mu_{B\cap C}(\omega), \nu_{B\cap C})$$

are smaller than

$$h_{\text{IFS}}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega)).$$

1. $\omega \in P_1 \cap Q_1$; by hypothesis, we have that:

   $$\max(\mu_A(\omega), \mu_B(\omega)) \leq \mu_C(\omega) \text{ and } \max(\nu_A(\omega), \nu_B(\omega)) \leq \nu_C(\omega),$$

   whence

   $$\begin{aligned} \mu_{A\cup C}(\omega) &= \mu_C(\omega), & \nu_{A\cup C}(\omega) &= \nu_A(\omega), \\ \mu_{A\cap C}(\omega) &= \mu_A(\omega), & \nu_{A\cap C}(\omega) &= \nu_C(\omega), \\ \mu_{B\cup C}(\omega) &= \mu_C(\omega), & \nu_{B\cup C}(\omega) &= \nu_B(\omega), \\ \mu_{B\cap C}(\omega) &= \mu_B(\omega), & \nu_{B\cap C}(\omega) &= \nu_C(\omega). \end{aligned}$$

   Moreover, property IF-loc.5 can be applied since

   $$\max(\nu_A(\omega), \nu_B(\omega)) + \mu_C(\omega) \leq \nu_C(\omega) + \mu_C(\omega) \leq 1,$$

   whence $(\nu_A(\omega), \mu_C(\omega)), (\nu_B(\omega), \mu_C(\omega)) \in T$ and therefore

   $$\begin{aligned} &h_{\text{IFS}}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega)) \\ &\quad \geq h_{\text{IFS}}(\mu_C(\omega), \nu_A(\omega), \mu_C(\omega), \nu_B(\omega)) \\ &\quad = h_{\text{IFS}}(\mu_{A\cup C}(\omega), \nu_{A\cup C}(\omega), \mu_{B\cup C}(\omega), \nu_{B\cup C}(\omega)). \end{aligned}$$

   Similarly,

   $$\begin{aligned} &h_{\text{IFS}}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega)) \\ &\quad \geq h_{\text{IFS}}(\mu_A(\omega), \nu_C(\omega), \mu_B(\omega), \nu_C(\omega)) \\ &\quad = h_{\text{IFS}}(\mu_{A\cap C}(\omega), \nu_{A\cap C}(\omega), \mu_{B\cap C}(\omega), \nu_{B\cap C}(\omega)). \end{aligned}$$

   Let us remark that, in the rest of the proof, axioms IF-loc.3, IF-loc.4 and IF-loc.5 are applicable because the previous hypotheses are satisfied.

2. $\omega \in P_1 \cap Q_2$; by hypothesis it holds that:

   $$\mu_A(\omega), \mu_B(\omega) \leq \mu_C(\omega) \text{ and } \nu_A(\omega) \leq \nu_C(\omega) < \nu_B(\omega),$$

   whence

   $$\begin{aligned} \mu_{A\cup C}(\omega) &= \mu_C(\omega), & \nu_{A\cup C}(\omega) &= \nu_A(\omega), \\ \mu_{A\cap C}(\omega) &= \mu_A(\omega), & \nu_{A\cap C}(\omega) &= \nu_C(\omega), \\ \mu_{B\cup C}(\omega) &= \mu_C(\omega), & \nu_{B\cup C}(\omega) &= \nu_C(\omega), \\ \mu_{B\cap C}(\omega) &= \mu_B(\omega), & \nu_{B\cap C}(\omega) &= \nu_B(\omega), \end{aligned}$$

As a consequence, by IF-loc.4 and IF-lo c.5:

$$\begin{aligned} &h_{IFS}(\mu_A(\omega),\nu_A(\omega),\mu_B(\omega),\nu_B(\omega)) \\ &\quad \geq h_{IFS}(\mu_A(\omega),\nu_A(\omega),\mu_B(\omega),\nu_C(\omega)) \\ &\quad \geq h_{IFS}(\mu_C(\omega),\nu_A(\omega),\mu_C(\omega),\nu_C(\omega)) \\ &\quad = h_{IFS}(\mu_{A\cup C}(\omega),\nu_{A\cup C}(\omega),\mu_{B\cup C}(\omega),\nu_{B\cup C}(\omega)). \end{aligned}$$

Similarly, by IF-lo c.4:

$$\begin{aligned} &h_{IFS}(\mu_A(\omega),\nu_A(\omega),\mu_B(\omega),\nu_B(\omega)) \\ &\quad \geq h_{IFS}(\mu_A(\omega),\nu_C(\omega),\mu_B(\omega),\nu_B(\omega)) \\ &\quad = h_{IFS}(\mu_{A\cap C}(\omega),\nu_{A\cap C}(\omega),\mu_{B\cap C}(\omega),\nu_{B\cap C}(\omega)). \end{aligned}$$

3. $\omega \in P_1 \cap Q_3$; this case is immediate from case 2, if we exchange the roles of A and B.

4. $\omega \in P_1 \cap Q_4$; then we know that:

$$\mu_A(\omega),\mu_B(\omega) \leq \mu_C(\omega), \text{ and } \nu_C(\omega) < \nu_A(\omega),\nu_B(\omega).$$

Then, it holds that $A \cup C = B \cup C = C$, $A \cap C = A$ and $B \cap C = B$, whence

$$h_{IFS}(\mu_A(\omega),\nu_A(\omega),\mu_B(\omega),\nu_B(\omega)) = h_{IFS}(\mu_{A\cap C}(\omega),\nu_{A\cap C}(\omega),\mu_{B\cap C}(\omega),\nu_{B\cap C}(\omega)).$$

Moreover,

$$\begin{aligned} h_{IFS}(\mu_A(\omega),\nu_A(\omega),\mu_B(\omega),\nu_B(\omega)) &\geq 0 = h_{IFS}(\mu_C(\omega),\nu_C(\omega),\mu_C(\omega),\nu_C(\omega)) \\ &= h_{IFS}(\mu_{A\cup C}(\omega),\nu_{A\cup C}(\omega),\mu_{B\cup C}(\omega),\nu_{B\cup C}(\omega)). \end{aligned}$$

5. $\omega \in P_2 \cap Q_1$; in that case we know that:

$$\mu_A(\omega) \leq \mu_C(\omega) < \mu_B(\omega) \text{ and } \nu_A(\omega),\nu_B(\omega) \leq \nu_C(\omega),$$

whence

$$\begin{aligned} \mu_{A\cup C}(\omega) &= \mu_C(\omega), & \nu_{A\cup C}(\omega) &= \nu_A(\omega), \\ \mu_{A\cap C}(\omega) &= \mu_A(\omega), & \nu_{A\cap C}(\omega) &= \nu_C(\omega), \\ \mu_{B\cup C}(\omega) &= \mu_B(\omega), & \nu_{B\cup C}(\omega) &= \nu_B(\omega), \\ \mu_{B\cap C}(\omega) &= \mu_C(\omega), & \nu_{B\cap C}(\omega) &= \nu_C(\omega). \end{aligned}$$

Thus, by IF-lo c.3,

$$\begin{aligned} &h_{IFS}(\mu_A(\omega),\nu_A(\omega),\mu_B(\omega),\nu_B(\omega)) \\ &\quad \geq h_{IFS}(\mu_C(\omega),\nu_A(\omega),\mu_B(\omega),\nu_B(\omega)) \\ &\quad = h_{IFS}(\mu_{A\cup C}(\omega),\nu_{A\cup C}(\omega),\mu_{B\cup C}(\omega),\nu_{B\cup C}(\omega)). \end{aligned}$$

Similarly, by IF-lo c.1 and IF-lo c.3,

$$\begin{aligned} &h_{IFS}(\mu_A(\omega),\nu_A(\omega),\mu_B(\omega),\nu_B(\omega)) \\ &\quad \geq 0 = h_{IFS}(\mu_C(\omega),\nu_C(\omega),\mu_C(\omega),\nu_C(\omega)) \\ &\quad \geq h_{IFS}(\mu_A(\omega),\nu_C(\omega),\mu_C(\omega),\nu_C(\omega)) \\ &\quad = h_{IFS}(\mu_{A\cap C}(\omega),\nu_{A\cap C}(\omega),\mu_{B\cap C}(\omega),\nu_{B\cap C}(\omega)). \end{aligned}$$

6. $\omega \quad P_2 \cap Q_2$; we know that:

$$\mu_A(\omega) \leq \mu_C(\omega) < \mu_B(\omega) \text{ and } \nu_A(\omega) \leq \nu_C(\omega) < \nu_B(\omega).$$

Then

$$\begin{aligned}
&\mu_{A\ \ C}(\omega) = \mu_C(\omega), \quad \nu_{A\ \ C}(\omega) = \nu_A(\omega),\\
&\mu_{A\cap C}(\omega) = \mu_A(\omega), \quad \nu_{A\cap C}(\omega) = \nu_C(\omega),\\
&\mu_{B\ \ C}(\omega) = \mu_B(\omega), \quad \nu_{B\ \ C}(\omega) = \nu_C(\omega),\\
&\mu_{B\cap C}(\omega) = \mu_C(\omega), \quad \nu_{B\cap C}(\omega) = \nu_B(\omega),
\end{aligned}$$

and therefore, by IF-loc.3 and IF-lo c.4,

$$\begin{aligned}
&h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad \geq h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_C(\omega))\\
&\quad \geq h_{IFS}(\mu_C(\omega), \nu_A(\omega), \mu_B(\omega), \nu_C(\omega))\\
&\quad = h_{IFS}(\mu_{A\ \ C}(\omega), \nu_{A\ \ C}(\omega), \mu_{B\ \ C}(\omega), \nu_{B\ \ C}(\omega)).
\end{aligned}$$

As a consequence,

$$\begin{aligned}
&h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad \geq h_{IFS}(\mu_A(\omega), \nu_C(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad \geq h_{IFS}(\mu_A(\omega), \nu_C(\omega), \mu_C(\omega), \nu_B(\omega))\\
&\quad = h_{IFS}(\mu_{A\cap C}(\omega), \nu_{A\cap C}(\omega), \mu_{B\cap C}(\omega), \nu_{B\cap C}(\omega)).
\end{aligned}$$

7. $\omega \quad P_2 \cap Q_3$; we know that:

$$\mu_A(\omega) \leq \mu_C(\omega) < \mu_B(\omega) \text{ and } \nu_B(\omega) \leq \nu_C(\omega) < \nu_A(\omega).$$

Thus,

$$\begin{aligned}
&\mu_{A\ \ C}(\omega) = \mu_C(\omega), \quad \nu_{A\ \ C}(\omega) = \nu_C(\omega),\\
&\mu_{A\cap C}(\omega) = \mu_A(\omega), \quad \nu_{A\cap C}(\omega) = \nu_A(\omega),\\
&\mu_{B\ \ C}(\omega) = \mu_B(\omega), \quad \nu_{B\ \ C}(\omega) = \nu_B(\omega),\\
&\mu_{B\cap C}(\omega) = \mu_C(\omega), \quad \nu_{B\cap C}(\omega) = \nu_C(\omega)
\end{aligned}$$

whence, applying IF-loc.3 and IF-lo c.4,

$$\begin{aligned}
&h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad \geq h_{IFS}(\mu_A(\omega), \nu_C(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad \geq h_{IFS}(\mu_C(\omega), \nu_C(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad = h_{IFS}(\mu_{A\ \ C}(\omega), \nu_{A\ \ C}(\omega), \mu_{B\ \ C}(\omega), \nu_{B\ \ C}(\omega)),
\end{aligned}$$

and as aconsequence

$$\begin{aligned}
&h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad \geq h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_C(\omega), \nu_B(\omega))\\
&\quad \geq h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_C(\omega), \nu_C(\omega))\\
&\quad = h_{IFS}(\mu_{A\cap C}(\omega), \nu_{A\cap C}(\omega), \mu_{B\cap C}(\omega), \nu_{B\cap C}(\omega)).
\end{aligned}$$

8. $\omega \in P_2 \cap Q_4$; it holds that:

$$\mu_A(\omega) \leq \mu_C(\omega) < \mu_B(\omega) \text{ and } \nu_C(\omega) \leq \nu_A(\omega), \nu_B(\omega),$$

whence

$$\begin{aligned}
\mu_{A \cup C}(\omega) &= \mu_C(\omega), & \nu_{A \cup C}(\omega) &= \nu_C(\omega),\\
\mu_{A \cap C}(\omega) &= \mu_A(\omega), & \nu_{A \cap C}(\omega) &= \nu_A(\omega),\\
\mu_{B \cup C}(\omega) &= \mu_B(\omega), & \nu_{B \cup C}(\omega) &= \nu_C(\omega),\\
\mu_{B \cap C}(\omega) &= \mu_C(\omega), & \nu_{B \cap C}(\omega) &= \nu_B(\omega).
\end{aligned}$$

and thus, by IF-lo c.3,

$$\begin{aligned}
&h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad 0 = h_{IFS}(\mu_C(\omega), \nu_C(\omega), \mu_C(\omega), \nu_C(\omega))\\
&\quad \geq h_{IFS}(\mu_C(\omega), \nu_C(\omega), \mu_B(\omega), \nu_C(\omega))\\
&\quad = h_{IFS}(\mu_{A \cup C}(\omega), \nu_{A \cup C}(\omega), \mu_{B \cup C}(\omega), \nu_{B \cup C}(\omega)).
\end{aligned}$$

In addition,

$$\begin{aligned}
&h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad \geq h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_C(\omega), \nu_B(\omega))\\
&\quad = h_{IFS}(\mu_{A \cap C}(\omega), \nu_{A \cap C}(\omega), \mu_{B \cap C}(\omega), \nu_{B \cap C}(\omega)).
\end{aligned}$$

9. $\omega \in P_3 \cap Q_i$; this case is immediate if we exchange the roles of A and B and apply the case when $\omega \in P_2 \cap Q_i$.

10. $\omega \in P_4 \cap Q_1$; in such case

$$\mu_C(\omega) < \mu_A(\omega), \mu_B(\omega) \text{ and } \nu_A(\omega), \nu_B(\omega) \leq \nu_C(\omega).$$

We have that:

$$\begin{aligned}
\mu_{A \cup C}(\omega) &= \mu_A(\omega), & \nu_{A \cup C}(\omega) &= \nu_A(\omega),\\
\mu_{A \cap C}(\omega) &= \mu_C(\omega), & \nu_{A \cap C}(\omega) &= \nu_C(\omega),\\
\mu_{B \cup C}(\omega) &= \mu_B(\omega), & \nu_{B \cup C}(\omega) &= \nu_B(\omega),\\
\mu_{B \cap C}(\omega) &= \mu_C(\omega), & \nu_{B \cap C}(\omega) &= \nu_C(\omega),
\end{aligned}$$

whence

$$h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega)) = h_{IFS}(\mu_{A \cup C}(\omega), \nu_{A \cup C}(\omega), \mu_{B \cup C}(\omega), \nu_{B \cup C}(\omega)),$$

and moreover, by IF-lo c.1,

$$\begin{aligned}
&h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad = 0 \geq h_{IFS}(\mu_C(\omega), \nu_C(\omega), \mu_C(\omega), \nu_C(\omega))\\
&\quad = h_{IFS}(\mu_{A \cap C}(\omega), \nu_{A \cap C}(\omega), \mu_{B \cap C}(\omega), \nu_{B \cap C}(\omega)).
\end{aligned}$$

11. $\omega \in P_4 \cap Q_2$; in such case we know that

$$\mu_C(\omega) < \mu_A(\omega), \mu_B(\omega) \text{ and } \nu_A(\omega) \leq \nu_C(\omega) < \nu_B(\omega).$$

It holds that

$$\begin{aligned}
\mu_{A\cup C}(\omega) &= \mu_A(\omega), & \nu_{A\cup C}(\omega) &= \nu_A(\omega),\\
\mu_{A\cap C}(\omega) &= \mu_C(\omega), & \nu_{A\cap C}(\omega) &= \nu_C(\omega),\\
\mu_{B\cup C}(\omega) &= \mu_B(\omega), & \nu_{B\cup C}(\omega) &= \nu_C(\omega),\\
\mu_{B\cap C}(\omega) &= \mu_C(\omega), & \nu_{B\cap C}(\omega) &= \nu_B(\omega),
\end{aligned}$$

whence, applying IF-loc.4,

$$\begin{aligned}
&h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad \geq h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_C(\omega))\\
&\quad = h_{IFS}(\mu_{A\cup C}(\omega), \nu_{A\cup C}(\omega), \mu_{B\cup C}(\omega), \nu_{B\cup C}(\omega)).
\end{aligned}$$

Moreover, applying IF-loc.1 and IF-loc.4,

$$\begin{aligned}
&h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad = 0 \geq h_{IFS}(\mu_C(\omega), \nu_C(\omega), \mu_C(\omega), \nu_C(\omega))\\
&\quad \geq h_{IFS}(\mu_C(\omega), \nu_C(\omega), \mu_C(\omega), \nu_B(\omega))\\
&\quad = h_{IFS}(\mu_{A\cap C}(\omega), \nu_{A\cap C}(\omega), \mu_{B\cap C}(\omega), \nu_{B\cap C}(\omega)).
\end{aligned}$$

12. $\omega \in P_4 \cap Q_3$; this follows from the previous case by exchanging the roles of A and B.

13. $\omega \in P_4 \cap Q_4$; we know that

$$\mu_C(\omega) < \mu_A(\omega), \mu_B(\omega) \text{ and } \nu_C(\omega) < \nu_A(\omega), \nu_B(\omega)$$

whence

$$\begin{aligned}
\mu_{A\cup C}(\omega) &= \mu_A(\omega), & \nu_{A\cup C}(\omega) &= \nu_C(\omega),\\
\mu_{A\cap C}(\omega) &= \mu_C(\omega), & \nu_{A\cap C}(\omega) &= \nu_A(\omega),\\
\mu_{B\cup C}(\omega) &= \mu_B(\omega), & \nu_{B\cup C}(\omega) &= \nu_C(\omega),\\
\mu_{B\cap C}(\omega) &= \mu_C(\omega), & \nu_{B\cap C}(\omega) &= \nu_B(\omega),
\end{aligned}$$

and thus by IF-loc.5

$$\begin{aligned}
&h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad \geq h_{IFS}(\mu_A(\omega), \nu_C(\omega), \mu_B(\omega), \nu_C(\omega))\\
&\quad = h_{IFS}(\mu_{A\cap C}(\omega), \nu_{A\cap C}(\omega), \mu_{B\cap C}(\omega), \nu_{B\cap C}(\omega)).
\end{aligned}$$

Moreover:

$$\begin{aligned}
&h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega))\\
&\quad \geq h_{IFS}(\mu_C(\omega), \nu_A(\omega), \mu_C(\omega), \nu_B(\omega))\\
&\quad = h_{IFS}(\mu_{A\cap C}(\omega), \nu_{A\cap C}(\omega), \mu_{B\cap C}(\omega), \nu_{B\cap C}(\omega)).
\end{aligned}$$

Hence, since $\Omega = \bigcup_{i=1}^{4} \bigcup_{j=1}^{4} P_i \cap Q_j$, we conclude that for all $\omega \in \Omega$ it holds that:

$$\begin{aligned} h_{\text{IFS}}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega)) \geq \\ \max\{ h_{\text{IFS}}(\mu_{A \cup C}(\omega), \nu_{A \cup C}(\omega), \mu_{B \cup C}(\omega), \nu_{B \cup C}(\omega)), \\ h_{\text{IFS}}(\mu_{A \cap C}(\omega), \nu_{A \cap C}(\omega), \mu_{B \cap C}(\omega), \nu_{B \cap C}(\omega))\}. \end{aligned}$$

Thus, D_{IFS} satisfies both IF-Div.3 and IF-Div.4, and therefore it is an IF-divergence. It only remains to show that D_{IFS} is local. But this holds trivially, taking into account that

$$\begin{aligned} & D_{\text{IFS}}(A, B) - D_{\text{IFS}}(A \cup \{\omega_j\}, B \cup \{\omega_j\}) \\ = & \sum_{i=1}^{n} h_{\text{IFS}}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)) \\ - & \sum_{i \neq j} h_{\text{IFS}}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)) - h_{\text{IFS}}(1, 1, 0, 0) \\ = & h_{\text{IFS}}(\mu_A(\omega_j), \nu_A(\omega_j), \mu_B(\omega_j), \nu_B(\omega_j)). \end{aligned}$$

We conclude that D_{IFS} is a local IF-divergence. ∎

Properties of local IF-divergences

In this section we are going to study some properties of local IF-divergences. In some cases, the local property will allows us to obtain interesting and useful properties.

We begin by studying under which conditions a local divergence satisfies IF-Div.5.

Proposition 5.30 *Let D_{IFS} be a local IF-divergence which associated function h_{IFS}. It satisfies IF-Div.5 if and only if for every*

$$(x_1, x_2), (y_1, y_2) \in T = \{(x, y) \in [0, 1]^2 \mid x + y \leq 1\}$$

it holds that

$$h_{\text{IFS}}(x_1, x_2, y_1, y_2) = h_{\text{IFS}}(x_2, x_1, y_2, y_1).$$

Proof Assume that D_{IFS} satisfies axiom IF-Div.5, i.e., that for every $A, B \in IFSs(\Omega)$, $D_{\text{IFS}}(A, B) = D_{\text{IFS}}(A^c, B^c)$. Consider $(x_1, x_2), (y_1, y_2) \in T$, and define the IF-sets A and B by:

$$A = \{(\omega, x_1, x_2) \mid \omega \in \Omega\} \text{ and } B = \{(\omega, y_1, y_2) \mid \omega \in \Omega\}.$$

By IF-Div.5, it holds that $D_{\text{IFS}}(A, B) = D_{\text{IFS}}(A^c, B^c)$. Using Equation (5.4),

$$n \, h_{\text{IFS}}(x_1, x_2, y_1, y_2) = D_{\text{IFS}}(A, B) = D_{\text{IFS}}(A^c, B^c) = n \, h_{\text{IFS}}(x_2, x_1, y_2, y_1).$$

Thus, $h_{\text{IFS}}(x_1, x_2, y_1, y_2) = h_{\text{IFS}}(x_2, x_1, y_2, y_1)$.

Converse ly, assume that $h_{IFS}(x_1, x_2, y_1, y_2) = h_{IFS}(x_2, x_1, y_2, y_1)$ for every two elements $(x_1, x_2), (y_1, y_2) \in T$. Let A and B be two IF-sets. Then, for every $i = 1, \ldots, n$ it holds that:

$$h_{IFS}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)) = h_{IFS}(\nu_A(\omega_i), \mu_A(\omega_i), \nu_B(\omega_i), \mu_B(\omega_i))$$

and therefore $D_{IFS}(A, B) = D_{IFS}(A^c, B^c)$. ∎

Next we give a le mma that shall b e useful later.

Lemma 5.31 *If D_{IFS} is a local IF-divergence, then for every $i = 1, \ldots, n$ it holds that*

$$D_{IFS}(A \cup \{\omega_i\}, B \cup \{\omega_i\}) = D_{IFS}(A \cap \{\omega_i\}^c, B \cap \{\omega_i\}^c).$$

Pro of Consider the IF-sets $A \cap \{\omega_i\}^c$ and $B \cap \{\omega_i\}^c$. Note that

$$(A \cap \{\omega_i\}^c) \cup \{\omega_i\} = (A \cup \{\omega_i\}) \cap (\{\omega_i\}^c \cup \{\omega_i\}) = A \cup \{\omega_i\}.$$
$$(B \cap \{\omega_i\}^c) \cup \{\omega_i\} = (B \cup \{\omega_i\}) \cap (\{\omega_i\}^c \cup \{\omega_i\}) = B \cup \{\omega_i\}.$$

Since D_{IFS} is a lo cal IF-divergence,

$$\begin{aligned} &D_{IFS}\left(A \cap \{\omega_i\}^c, B \cap \{\omega_i\}^c\right) - D_{IFS}\left((A \cap \{\omega_i\}^c) \cup \{\omega_i\}, (B \cap \{\omega_i\}^c) \cup \{\omega_i\}\right) \\ &= D_{IFS}(A \cap \{\omega_i\}^c, B \cap \{\omega_i\}^c) - D_{IFS}(A \cup \{\omega_i\}, B \cup \{\omega_i\}) \\ &= h_{IFS}(\mu_{A \cap \{\omega_i\}^c}(\omega_i), \nu_{A \cap \{\omega_i\}^c}(\omega_i), \mu_{B \cap \{\omega_i\}^c}(\omega_i), \nu_{B \cap \{\omega_i\}^c}(\omega_i)) \\ &= h_{IFS}(0, 1, 0, 1) = 0, \end{aligned}$$

using that

$$\begin{aligned} \mu_{A \cap \{\omega_i\}^c}(\omega_i) &= \min(\mu_A(\omega_i), 0) = 0, \\ \nu_{A \cap \{\omega_i\}^c}(\omega_i) &= \max(\nu_A(\omega_i), 1) = 1, \\ \mu_{B \cap \{\omega_i\}^c}(\omega_i) &= \min(\mu_B(\omega_i), 0) = 0, \\ \nu_{B \cap \{\omega_i\}^c}(\omega_i) &= \max(\mu_B(\omega_i), 1) = 1. \end{aligned}$$

Using this lemma, we can establish the following prop osition.

Prop osition 5.32 *An IF-divergence D_{IFS} is local if and only if there is a function h such that*

$$D_{IFS}(A, B) - D_{IFS}(A \cap \{\omega_i\}^c, B \cap \{\omega_i\}^c) = h(\mu_A(\omega_i), \nu_A(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i))$$

for every $A, B \in IFSs(\Omega)$.

Pro of It is immediate from the previous lemma. ∎

Let us give another characterization of lo cal IF-divergences.

Proposition 5.33 *An IF-divergence D_{IFS} is local if and only if for every $X \in \mathcal{P}(\Omega)$ it holds that:*

$$D_{IFS}(A,B) = D_{IFS}(A \cap X, B \cap X) + D_{IFS}(A \cap X^c, B \cap X^c),$$

for every $A,B \in IFSs(\Omega)$.

Proof Assume that D_{IFS} is a local IF-divergence, and let us consider $A,B \in IFSs(\Omega)$ and $X \in \mathcal{P}(\Omega)$.

Since $A = (A \cap X) \cup (A \cap X^c)$ and $B = (B \cap X) \cup (B \cap X^c)$, it holds that

$$D_{IFS}(A,B) = D_{IFS}((A \cap X) \cup (A \cap X^c), (B \cap X) \cup (B \cap X^c)).$$

Taking into account that D_{IFS} is local, we deduce that:

$$\begin{aligned} D_{IFS}(A,B) = \sum_{i=1}^{n} h_{IFS}(&\mu_{(A \cap X) \cup (A \cap X^c)}(\omega_i), \nu_{(A \cap X) \cup (A \cap X^c)}(\omega_i),\\ &\mu_{(B \cap X) \cup (B \cap X^c)}(\omega_i), \nu_{(B \cap X) \cup (B \cap X^c)}(\omega_i)). \end{aligned}$$

Moreover, by splitting the sum between the elements on X and X^c,

$$\begin{aligned} D_{IFS}(A,B) = &\sum_{\omega \in X} h_{IFS}(\mu_{(A \cap X) \cup (A \cap X^c)}(\omega), \nu_{(A \cap X) \cup (A \cap X^c)}(\omega),\\ &\qquad \mu_{(B \cap X) \cup (B \cap X^c)}(\omega), \nu_{(B \cap X) \cup (B \cap X^c)}(\omega))\\ + &\sum_{\omega \in X^c} h_{IFS}(\mu_{(A \cap X) \cup (A \cap X^c)}(\omega), \nu_{(A \cap X) \cup (A \cap X^c)}(\omega),\\ &\qquad \mu_{(B \cap X) \cup (B \cap X^c)}(\omega), \nu_{(B \cap X) \cup (B \cap X^c)}(\omega)). \end{aligned}$$

Furthermore:

$$\omega \in X \Rightarrow \begin{cases} \mu_{(A \cap X) \cup (A \cap X^c)}(\omega) = \max(\mu_{A \cap X}(\omega), \mu_{A \cap X^c}(\omega)) \\ \qquad = \max(\mu_{A \cap X}(\omega), 0) = \mu_{A \cap X}(\omega). \\ \nu_{(A \cap X) \cup (A \cap X^c)}(\omega) = \min(\nu_{A \cap X}(\omega), \nu_{A \cap X^c}(\omega)) \\ \qquad = \min(\nu_{A \cap X}(\omega), 1) = \nu_{A \cap X}(\omega). \end{cases}$$

$$\omega \in X^c \Rightarrow \begin{cases} \mu_{(A \cap X) \cup (A \cap X^c)}(\omega) = \max(\mu_{A \cap X}(\omega), \mu_{A \cap X^c}(\omega)) \\ \qquad = \max(0, \mu_{A \cap X^c}(\omega)) = \mu_{A \cap X^c}(\omega). \\ \nu_{(A \cap X) \cup (A \cap X^c)}(\omega) = \min(\nu_{A \cap X}(\omega), \nu_{A \cap X^c}(\omega)) \\ \qquad = \min(1, \nu_{A \cap X^c}(\omega)) = \nu_{A \cap X^c}(\omega). \end{cases}$$

Similarly,

$$\omega \in X \Rightarrow \begin{array}{l} \mu_{(B \cap X) \cup (B \cap X^c)}(\omega) = \mu_{B \cap X}(\omega). \\ \nu_{(B \cap X) \cup (B \cap X^c)}(\omega) = \nu_{B \cap X}(\omega). \end{array}$$

$$\omega \in X^c \Rightarrow \begin{array}{l} \mu_{(B \cap X) \cup (B \cap X^c)}(\omega) = \mu_{B \cap X^c}(\omega). \\ \nu_{(B \cap X) \cup (B \cap X^c)}(\omega) = \nu_{B \cap X^c}(\omega). \end{array}$$

Thus, the expression of $D_{\mathrm{IFS}}(A,B)$ becomes

$$D_{\mathrm{IFS}}(A,B)=\sum_{\omega\in X} h_{\mathrm{IFS}}(\mu_{A\cap X}(\omega),\nu_{A\cap X}(\omega),\mu_{B\cap X}(\omega),\nu_{B\cap X}(\omega))$$
$$+\sum_{\omega\in X^c} h_{\mathrm{IFS}}(\mu_{A\cap X^c}(\omega),\nu_{A\cap X^c}(\omega),\mu_{B\cap X^c}(\omega),\nu_{B\cap X^c}(\omega)).$$

Taking into account that

$$\begin{aligned}
D_{\mathrm{IFS}}(A\cap X,B\cap X) &= \sum_{\omega\in\Omega} h_{\mathrm{IFS}}(\mu_{A\cap X}(\omega),\nu_{A\cap X}(\omega),\mu_{B\cap X}(\omega),\nu_{B\cap X}(\omega))\\
&= \sum_{\omega\in X} h_{\mathrm{IFS}}(\mu_{A\cap X}(\omega),\nu_{A\cap X}(\omega),\mu_{B\cap X}(\omega),\nu_{B\cap X}(\omega))\\
&\quad+\sum_{\omega\in X^c} h_{\mathrm{IFS}}(\mu_{A\cap X}(\omega),\nu_{A\cap X}(\omega),\mu_{B\cap X}(\omega),\nu_{B\cap X}(\omega))\\
&= \sum_{\omega\in X} h_{\mathrm{IFS}}(\mu_{A\cap X}(\omega),\nu_{A\cap X}(\omega),\mu_{B\cap X}(\omega),\nu_{B\cap X}(\omega)),\\
D_{\mathrm{IFS}}(A\cap X^c,B\cap X^c) &= \sum_{\omega\in\Omega} h_{\mathrm{IFS}}(\mu_{A\cap X^c}(\omega),\nu_{A\cap X^c}(\omega),\mu_{B\cap X^c}(\omega),\nu_{B\cap X^c}(\omega))\\
&= \sum_{\omega\in X} h_{\mathrm{IFS}}(\mu_{A\cap X^c}(\omega),\nu_{A\cap X^c}(\omega),\mu_{B\cap X^c}(\omega),\nu_{B\cap X^c}(\omega))\\
&\quad+\sum_{\omega\in X^c} h_{\mathrm{IFS}}(\mu_{A\cap X^c}(\omega),\nu_{A\cap X^c}(\omega),\mu_{B\cap X^c}(\omega),\nu_{B\cap X^c}(\omega))\\
&= \sum_{\omega\in X^c} h_{\mathrm{IFS}}(\mu_{A\cap X^c}(\omega),\nu_{A\cap X^c}(\omega),\mu_{B\cap X^c}(\omega),\nu_{B\cap X^c}(\omega)),
\end{aligned}$$

we conclude that

$$D_{\mathrm{IFS}}(A,B)=D_{\mathrm{IFS}}(A\cap X,B\cap X)+D_{\mathrm{IFS}}(A\cap X^c,B\cap X^c).$$

Conversely, assume that $D_{\mathrm{IFS}}(A,B)=D_{\mathrm{IFS}}(A\cap X,B\cap X)+D_{\mathrm{IFS}}(A\cap X^c,B\cap X^c)$ for every $A,B\in IFSs(\Omega)$ and $X\subseteq\Omega$. Applying this property to the crisp set $X=\{\omega_1\}$,

$$\begin{aligned}
D_{\mathrm{IFS}}(A,B) &= D_{\mathrm{IFS}}(A\cap\{\omega_1\},B\cap\{\omega_1\})+D_{\mathrm{IFS}}(A\cap\{\omega_2,\dots,\omega_n\},B\cap\{\omega_2,\dots,\omega_n\})\\
&= D_{\mathrm{IFS}}(A_1,B_1)+D_{\mathrm{IFS}}(A\cap\{\omega_2,\dots,\omega_n\},B\cap\{\omega_2,\dots,\omega_n\}),
\end{aligned}$$

where the IF-sets A_1 and B_1 are defined by

$$\begin{aligned}
A_1 &= \{(\omega_1,\mu_A(\omega_1),\nu_A(\omega_1)),(\omega_i,0,1)\mid i\neq 1\},\\
B_1 &= \{(\omega_1,\mu_B(\omega_1),\nu_B(\omega_1)),(\omega_i,0,1)\mid i\neq 1\}.
\end{aligned}$$

Now, apply the hypothesis to the crisp set $X=\{\omega_2\}$ and the IF-sets $A\cap\{\omega_2,\dots,\omega_n\}$ and $B\cap\{\omega_2,\dots,\omega_n\}$.

$$\begin{aligned}
D_{\mathrm{IFS}}(A\cap\{\omega_2,\dots,\omega_n\},B\cap\{\omega_2,\dots,\omega_n\}) &= D_{\mathrm{IFS}}(A\cap\{\omega_2\},B\cap\{\omega_2\})\\
&\quad+D_{\mathrm{IFS}}(A\cap\{\omega_3,\dots,\omega_n\},B\cap\{\omega_3,\dots,\omega_n\})\\
&= D_{\mathrm{IFS}}(A_2,B_2)\\
&\quad+D_{\mathrm{IFS}}(A\cap\{\omega_3,\dots,\omega_n\},B\cap\{\omega_3,\dots,\omega_n\}),
\end{aligned}$$

where

$$A_2 = \{(\omega_2, \mu_A(\omega_2), \nu_A(\omega_2))\} \cup \{(\omega_i, 0, 1) \mid i \neq 2\},$$
$$B_2 = \{(\omega_2, \mu_B(\omega_2), \nu_B(\omega_2))\} \cup \{(\omega_i, 0, 1) \mid i \neq 2\}.$$

If we repeat the process, for any $j \in \{1, \dots, n-1\}$, given $X = \{\omega_j\}$ and the IF-sets $A \cap \{\omega_j, \dots, \omega_n\}$ and $B \cap \{\omega_j, \dots, \omega_n\}$, it holds that:

$$\begin{aligned} D_{IFS}(A \cap \{\omega_j, \dots, \omega_n\}, B \cap \{\omega_j, \dots, \omega_n\}) \\ &= D_{IFS}(A \cap \{\omega_j\}, B \cap \{\omega_j\}) + D_{IFS}(A \cap \{\omega_{j+1}, \dots, \omega_n\}, B \cap \{\omega_{j+1}, \dots, \omega_n\}) \\ &= D_{IFS}(A_j, B_j) + D_{IFS}(A \cap \{\omega_{j+1}, \dots, \omega_n\}, B \cap \{\omega_{j+1}, \dots, \omega_n\}), \end{aligned}$$

where

$$A_j = \{(\omega_j, \mu_A(\omega_j), \nu_A(\omega_j))\} \cup \{(\omega_i, 0, 1) \mid i \neq j\},$$
$$B_j = \{(\omega_j, \mu_B(\omega_j), \nu_B(\omega_j))\} \cup \{(\omega_i, 0, 1) \mid i \neq j\}.$$

Then, $D_{IFS}(A, B)$ can be expressed by

$$\begin{aligned} D_{IFS}(A, B) &= D_{IFS}(A_1, B_1) + D_{IFS}(A \cap \{\omega_2, \dots, \omega_n\}, B \cap \{\omega_2, \dots, \omega_n\}) \\ &= D_{IFS}(A_1, B_1) + D_{IFS}(A_2, B_2) \\ &+ D_{IFS}(A \cap \{\omega_3, \dots, \omega_n\}, B \cap \{\omega_3, \dots, \omega_n\}) \\ &= \dots = \sum_{i=1}^{n} D_{IFS}(A_i, B_i). \end{aligned}$$

Now, consider the difference between $D_{IFS}(A, B)$ and $D_{IFS}(A \cap \{\omega_i\}, B \cap \{\omega_i\})$:

$$D_{IFS}(A \cup \{\omega_i\}, B \cup \{\omega_i\}) - D_{IFS}(A, B) = D_{IFS}(A_i \cup \{\omega_i\}, B_i \cup \{\omega_i\}) - D_{IFS}(A_i, B_i).$$

This difference only depends on $\mu_A(\omega_i), \nu_A(\omega_i)$ and $\mu_B(\omega_i), \nu_B(\omega_i)$, so taking into account Definition 5.28 we conclude that D_{IFS} is a local IF-divergence. ∎

A particular case of interest is the comparison of an IF-set and its complementary. In this sense, it seems useful to measure how imprecise an IF-set is. We consider the following partial order between IF-sets: given two IF-sets A and B, we say that A is sharper than B, and denote it $A \ll B$, when $|\mu_A(\omega) - 0.5| \geq |\mu_B(\omega) - 0.5|$ and $|\nu_A(\omega) - 0.5| \geq |\nu_B(\omega) - 0.5|$ for every $\omega \in \Omega$.

Using this partial order we can establish the following interesting property.

Proposition 5.34 *If D_{IFS} is a local IF-divergence and $A \ll B$, then it holds that $D_{IFS}(A, A^c) \geq D_{IFS}(B, B^c)$.*

Proof Assume that $A \ll B$, and let us consider the crisp sets X and Y defined by

$$X = \{\omega \in \Omega \mid \mu_A(\omega) \leq 0.5 \text{ and } \nu_A(\omega) \geq 0.5\}.$$
$$Y = \{\omega \in \Omega \mid \mu_B(\omega) \leq \nu_B(\omega)\}.$$

Applying Prop osition 5.33,

$$D_{IFS}(A,A^c) = D_{IFS}(A \cap X, A^c \cap X) + D_{IFS}(A \cap X^c, A^c \cap X^c)$$

and if we use the same prop osition with $D_{IFS}(A \cap X, A^c \cap X)$ and $D_{IFS}(A \cap X^c, A^c \cap X^c)$, we obtai n that

$$\begin{aligned} D_{IFS}(A \cap X, A^c \cap X) &= D_{IFS}(A \cap X \cap Y, A^c \cap X \cap Y) \\ &+ D_{IFS}(A \cap X \cap Y^c, A^c \cap X \cap Y^c), \\ D_{IFS}(A \cap X^c, A^c \cap X^c) &= D_{IFS}(A \cap X^c \cap Y, A^c \cap X^c \cap Y) \\ &+ D_{IFS}(A \cap X^c \cap Y^c, A^c \cap X^c \cap Y^c). \end{aligned}$$

Hence,

$$\begin{aligned} D_{IFS}(A,A^c) &= D_{IFS}(A \cap X \cap Y, A^c \cap X \cap Y) \\ &+ D_{IFS}(A \cap X \cap Y^c, A^c \cap X \cap Y^c) \\ &+ D_{IFS}(A \cap X^c \cap Y, A^c \cap X^c \cap Y) \\ &+ D_{IFS}(A \cap X^c \cap Y^c, A^c \cap X^c \cap Y^c). \end{aligned}$$

Letus studyeachof the summandsin the right-hand-sideseparately. For the firstone, we have that

$$\mu_{A \cap X \cap Y}(\omega) = \begin{cases} \mu_A(\omega) & \text{if } \mu_A(\omega) \leq 0.5 \leq \nu_A(\omega) \text{ and } \mu_B(\omega) \leq \nu_B(\omega), \\ 0 & \text{otherwise,} \end{cases}$$

$$\nu_{A \cap X \cap Y}(\omega) = \begin{cases} \nu_A(\omega) & \text{if } \mu_A(\omega) \leq 0.5 \leq \nu_A(\omega) \text{ and } \mu_B(\omega) \leq \nu_B(\omega), \\ 1 & \text{otherwise.} \end{cases}$$

Howe ver, if $\omega \quad X \cap Y$, taking into account that $A \ll B$, it holds that

$$\mu_A(\omega) \leq \mu_B(\omega) \leq 0.5 \leq \nu_B(\omega) \leq \nu_A(\omega)$$

and therefore,

$$A \cap X \cap Y \quad B \quad B^c \cap X \cap Y \quad A^c \cap X \cap Y.$$

Now, applying Lemma 5.5 we obtainthat

$$D_{IFS}(A \cap X \cap Y, A^c \cap X \cap Y) \geq D_{IFS}(B \cap X \cap Y, B^c \cap X \cap Y).$$

Letus considernext the second term

$$\mu_{A \cap X \cap Y^c}(\omega) = \begin{cases} \mu_A(\omega) & \text{if } \mu_A(\omega) \leq 0.5 \leq \nu_A(\omega) \text{ and } \nu_B(\omega) \leq \mu_B(\omega), \\ 0 & \text{otherwise,} \end{cases}$$

$$\nu_{A \cap X^c \cap Y}(\omega) = \begin{cases} \nu_A(\omega) & \text{if } \mu_A(\omega) \leq 0.5 \leq \nu_A(\omega) \text{ and } \nu_B(\omega) \leq \mu_B(\omega), \\ 1 & \text{otherwise,} \end{cases}$$

However, if $\omega \quad X \cap Y^c$, since $A \ll B$ itholds that

$$\mu_A(\omega) \leq \nu_B(\omega) \leq 0.5 \leq \mu_B(\omega) \leq \nu_A(\omega)$$

whence

$$A \cap X \cap Y^c \quad B^c \cap X \cap Y^c \quad B \cap X \cap Y^c \quad A^c \cap X \cap Y^c$$

and if we applyLemma 5.5 we obtain that

$$D_{\text{IFS}}(A \cap X \cap Y^c, A^c \cap X \cap Y^c) \geq D_{\text{IFS}}(B \cap X \cap Y^c, B^c \cap X \cap Y^c).$$

Considernext the third summand

$$\mu_{A \cap X^c \cap Y}(\omega) = \begin{cases} \mu_A(\omega) & \text{if } \nu_A(\omega) \leq 0.5 \leq \mu_A(\omega) \text{ and } \mu_B(\omega) \leq \nu_B(\omega), \\ 0 & \text{otherwise,} \end{cases}$$

$$\nu_{A \cap X^c \cap Y}(\omega) = \begin{cases} \nu_A(\omega) & \text{if } \nu_A(\omega) \leq 0.5 \leq \mu_A(\omega) \text{ and } \mu_B(\omega) \leq \nu_B(\omega), \\ 1 & \text{otherwise.} \end{cases}$$

If $\omega \quad X^c \cap Y$, since $A \ll B$, it holds that

$$\nu_A(\omega) \leq \mu_B(\omega) \leq 0.5 \leq \nu_B(\omega) \leq \mu_A(\omega)$$

whence

$$A^c \cap X^c \cap Y \quad B \cap X^c \cap Y \quad B^c \cap X^c \cap Y \quad A \cap X^c \cap Y.$$

Applying Lemma 5.5,we obtain that

$$D_{\text{IFS}}(A \cap X^c \cap Y, A^c \cap X^c \cap Y) \geq D_{\text{IFS}}(B \cap X^c \cap Y, B^c \cap X^c \cap Y)$$

Finally, consider the fourthterm:

$$\mu_{A \cap X^c \cap Y^c}(\omega) = \begin{cases} \mu_A(\omega) & \text{if } \nu_A(\omega) \leq 0.5 \leq \mu_A(\omega) \text{ and } \nu_B(\omega) \leq \mu_B(\omega), \\ 0 & \text{otherwise,} \end{cases}$$

$$\nu_{A \cap X^c \cap Y^c}(\omega) = \begin{cases} \nu_A(\omega) & \text{if } \nu_A(\omega) \leq 0.5 \leq \mu_A(\omega) \text{ and } \nu_B(\omega) \leq \mu_B(\omega), \\ 1 & \text{otherwise.} \end{cases}$$

If $\omega \quad X^c \cap Y^c$, taking into account that $A \ll B$, it hold s that:

$$\nu_A(\omega) \leq \nu_B(\omega) \leq 0.5 \leq \mu_B(\omega) \leq \mu_A(\omega).$$

Then, usin g Lemma 5.5 we obtain that

$$D_{\text{IFS}}(A \cap X^c \cap Y^c, A^c \cap X^c \cap Y^c) \geq D_{\text{IFS}}(B \cap X^c \cap Y^c, B^c \cap X^c \cap Y^c)$$

and therefore

$$\begin{aligned} D_{\text{IFS}}(A, A^c) = {} & D_{\text{IFS}}(A \cap X \cap Y, A^c \cap X \cap Y) \\ & + D_{\text{IFS}}(A \cap X \cap Y^c, A^c \cap X \cap Y^c) \\ & + D_{\text{IFS}}(A \cap X^c \cap Y, A^c \cap X^c \cap Y) \\ & + D_{\text{IFS}}(A \cap X^c \cap Y^c, A^c \cap X^c \cap Y^c) \\ \geq {} & D_{\text{IFS}}(B \cap X \cap Y, B^c \cap X \cap Y) \\ & + D_{\text{IFS}}(B \cap X \cap Y^c, B^c \cap X \cap Y^c) \\ & + D_{\text{IFS}}(B \cap X^c \cap Y, B^c \cap X^c \cap Y) \\ & + D_{\text{IFS}}(B \cap X^c \cap Y^c, B^c \cap X^c \cap Y^c) = D_{\text{IFS}}(B, B^c). \end{aligned}$$

This completes the proof. ∎

The above result implies that the lower the fuzziness, the greater the divergence between an IF-set and its complementary. Moreover, the divergence is maximum when the IF-set is crisp.

Proposition 5.35 *If V and Z are two crisp sets and D_{IFS} is a local IF-divergence,*

$$D_{IFS}(V, V^c) = D_{IFS}(Z, Z^c).$$

In addition, if $A, B \in IFSs(\Omega)$, then $D_{IFS}(A, B) \leq D_{IFS}(Z, Z^c)$.

Proof Note that, by IF-loc.2 of Theorem 5.29 $h_{IFS}(1, 0, 0, 1) = h_{IFS}(0, 1, 1, 0)$ and therefore

$$D_{IFS}(V, V^c) = n \quad h_{IFS}(1, 0, 0, 1) = D_{IFS}(Z, Z^c).$$

Now, taking into account that $h_{IFS}(1, 0, 0, 1) \geq h_{IFS}(x_1, x_2, y_1, y_2)$, since by IF-loc.3 and IF-loc.4:

$$\begin{aligned} h_{IFS}(1, 0, 0, 1) &\geq h_{IFS}(x_1, 0, 0, 1) \geq h_{IFS}(x_1, x_2, 0, 1) \\ &\geq h_{IFS}(x_1, x_2, 0, y_2) \geq h_{IFS}(x_1, x_2, y_1, y_2), \end{aligned}$$

we have that

$$\begin{aligned} D_{IFS}(A, B) &= \sum_{i=1}^{n} h_{IFS}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)) \\ &\leq \sum_{i=1}^{n} h_{IFS}(1, 0, 0, 1) = D_{IFS}(Z, Z^c). \end{aligned}$$

We have seen that every IF-divergence is also an IF-dissimilarity, and therefore it satisfies that $D_{IFS}(A, C) \geq \max(D_{IFS}(A, B), D_{IFS}(B, C))$ for every IF-sets A, B and C such that $A \subseteq B \subseteq C$. In the following proposition we obtain a similar result for local IF-divergences with less restrictive conditions.

Proposition 5.36 *Let D_{IFS} be a local IF-divergence. If for every $\omega \in \Omega$ either*

$$\mu_A(\omega) \leq \mu_B(\omega) \leq \mu_C(\omega) \text{ and } \nu_A(\omega) \geq \nu_B(\omega) \geq \nu_C(\omega),$$

or

$$\mu_A(\omega) \geq \mu_B(\omega) \geq \mu_C(\omega) \text{ and } \nu_A(\omega) \leq \nu_B(\omega) \leq \nu_C(\omega),$$

then $D_{IFS}(A, C) \geq \max(D_{IFS}(A, B), D_{IFS}(B, C))$.

Proof Since the IF-divergence is local we can apply properties IF-loc.3 and IF-loc.4, and we obtain the following:

$$\begin{aligned} D_{\mathrm{IFS}}(A, C) &= \sum_{i=1}^{n} h_{\mathrm{IFS}}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_C(\omega_i), \nu_C(\omega_i)) \\ &\geq \max\Big(\sum_{i=1}^{n} h_{\mathrm{IFS}}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)), \\ &\qquad \sum_{i=1}^{n} h_{\mathrm{IFS}}(\mu_B(\omega_i), \nu_B(\omega_i), \mu_C(\omega_i), \nu_C(\omega_i))\Big) \\ &= \max(D_{\mathrm{IFS}}(A, B), D_{\mathrm{IFS}}(B, C)). \end{aligned}$$

In Proposition 5.27 we proved that, if D_{IFS} is an IF-divergence, then $D^{\varphi}_{\mathrm{IFS}}$ is also an IF-divergence, where $D^{\varphi}_{\mathrm{IFS}}(A, B) = \varphi(D_{\mathrm{IFS}}(A, B))$ and φ is a increasing function such that $\varphi(0) = 0$. In particular, if D_{IFS} is a local IF-divergence, $D^{\varphi}_{\mathrm{IFS}}$ is local if and only if φ is linear. Next we derive a similar method to build local IF-divergences from local IF-divergences.

Proposition 5.37 *Let D_{IFS} be a local IF-divergence, and let $\varphi: [0, \infty) \to [0, \infty)$ be a increasing function such that $\varphi(0) = 0$. Then, the function $D_{\mathrm{IFS},\varphi}$, defined by*

$$D_{\mathrm{IFS},\varphi}(A, B) = \sum_{i=1}^{n} \varphi h_{\mathrm{IFS}}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)),$$

is a local IF-divergence.

Proof Immediate using the properties of φ and taking into account that h_{IFS} satisfies the properties IF-loc.1 to IF-loc.5. ■

To conclude this section, we relate local IF-divergences and real distances.

Proposition 5.38 *Consider a distance $d: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ satisfying*

$$\max(d(x, y), d(y, z)) \leq d(x, z)$$

for $x < y < z$. Then, for every increasing function $\varphi: [0, \infty) \times [0, \infty) \to [0, \infty)$ such that $\varphi(0, 0) = 0$, the function $D_{\mathrm{IFS}}: IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ defined by:

$$D_{\mathrm{IFS}}(A, B) = \sum_{i=1}^{n} \varphi(d(\mu_A(\omega_i), \mu_B(\omega_i)), d(\nu_A(\omega_i), \nu_B(\omega_i)))$$

is a local IF-divergence.

Proof Using Theorem 5.42, it suffices to prove that the function

$$h_{\mathrm{IFS}}(x_1, x_2, y_1, y_2) = \varphi(d(x_1, y_1), d(x_2, y_2))$$

satisfies the properties IF-loc.1 to IF-loc.5.

IF-loc.1: Consider $(x, y) \in T = \{(x, y) \in [0, 1]^2 \mid x+y \leq 1\}$. Since d is a distance, $d(x, x) = d(y, y) = 0$, and therefore

$$h_{\text{IFS}}(x, y, x, y) = \varphi(d(x, x), d(y, y)) = \varphi(0, 0) = 0.$$

IF-loc.2: Take (x_1, x_2) and (y_1, y_2) in T. Since d is a distance, $d(x_1, y_1) = d(y_1, x_1)$ and $d(x_2, y_2) = d(y_2, x_2)$, whence

$$h_{\text{IFS}}(x_1, x_2, y_1, y_2) = \varphi(d(x_1, y_1), d(x_2, y_2)) = \varphi(d(y_1, x_1), d(y_2, x_2)) = h_{\text{IFS}}(y_1, y_2, x_1, x_2).$$

IF-loc.3: Consider $(x_1, x_2), (y_1, y_2) \in T$ and $z \in [0, 1]$ such that $x_1 \leq z \leq y_1$. Applying the hypothesis on d,

$$d(x_1, y_1) \geq \max(d(x_1, z), d(z, y_1))$$

whence

$$h_{\text{IFS}}(x_1, x_2, y_1, y_2) = \varphi(d(x_1, y_1), d(x_2, y_2)) \geq \varphi(d(x_1, z), d(x_2, y_2)) = h_{\text{IFS}}(x_1, x_2, z, y_2).$$

Moreover, if $(x_2, z), (y_2, z) \in T$, then $\max(x_2, y_2) + z \leq 1$ and it holds that:

$$h_{\text{IFS}}(x_1, x_2, y_1, y_2) = \varphi(d(x_1, y_1), d(x_2, y_2)) \geq \varphi(d(z, y_1), d(x_2, y_2)) = h_{\text{IFS}}(z, x_2, y_1, y_2).$$

IF-loc.4: Let $(x_1, x_2), (y_1, y_2) \in T$ and $z \in [0, 1]$ such that $x_2 \leq z \leq y_2$. Applying the hypothesis on d,

$$d(x_2, y_2) \geq \max(d(x_2, z), d(z, y_2)).$$

Since φ is increasing in each component:

$$h_{\text{IFS}}(x_1, x_2, y_1, y_2) = \varphi(d(x_1, y_1), d(x_2, y_2)) \geq \varphi(d(x_1, y_1), d(x_2, z)) = h_{\text{IFS}}(x_1, x_2, y_1, z).$$

Moreover, if $(x_1, z), (y_1, z) \in T$, it holds that $\max(x_1, y_1) + z \leq 1$ and then:

$$h_{\text{IFS}}(x_1, x_2, y_1, y_2) = \varphi(d(x_1, y_1), d(x_2, y_2)) \geq \varphi(d(x_1, y_1), d(z, y_2)) = h_{\text{IFS}}(x_1, z, y_1, y_2).$$

IF-loc.5: Finally, consider $(x_1, x_2), (y_1, y_2) \in T$ and $z \in [0, 1]$. Applying our hypothesis on d, it holds that:

$$d(z, z) = 0 \leq \min(d(x_1, y_1), d(x_2, y_2)).$$

Then, if $(x_2, z), (y_2, z) \in T$, it holds that $\max(x_2, y_2) + z \leq 1$, and since φ is increasing in each component, it follows that

$$h_{\text{IFS}}(z, x_2, z, y_2) = \varphi(d(z, z), d(x_2, y_2)) \leq \varphi(d(x_1, y_1), d(x_2, y_2)) = h_{\text{IFS}}(x_1, x_2, y_1, y_2).$$

Moreover, if $(x_1, z), (x_2, z)$ T, then $\max(x_1, y_1) + z \leq 1$, and since φ is increasing in each comp onent, it holds that:

$$h_{\text{IFS}}(x_1, z, y_1, z) = \varphi(d(x_1, y_1), d(z, z)) \leq \varphi(d(x_1, y_1), d(x_2, y_2)) = h_{\text{IFS}}(x_1, x_2, y_1, y_2).$$

Thus, h_{IFS} satisfies properties IF-lo c.1 to IF-lo c.5. Applying Theorem 5.29, we conclude that D_{IFS} is a lo cal IF-divergence. ∎

Let us see an example of an application of this result.

Example 5.39 *Consider the distance d defined by $d(x, y) = |x - y|$, and the increasing function $\varphi(x, y) = \frac{x+y}{2n}$, that satisfies $\varphi(0, 0) = 0$. Then, we can define the function $D_{\text{IFS}} : IF\,Ss(\Omega) \times IF\,Ss(\Omega) \to \mathbb{R}$ defined by*

$$D_{\text{IFS}}(A, B) = \sum_{i=1}^{n} \varphi\left(d(\mu_A(\omega_i), \mu_B(\omega_i)), d(\nu_A(\omega_i), \nu_B(\omega_i))\right)$$

for every A, B $IF\,Ss(\Omega)$ is an IF-divergence. In fact, if we input the values of φ and d, D_{IFS} becomes

$$D_{\text{IFS}}(A, B) = \sum_{i=1}^{n} |\mu_A(\omega_i) - \mu_B(\omega_i)| + |\nu_A(\omega_i) - \nu_B(\omega_i)|,$$

i.e., we obtain Hong and Kim IF-divergence D_C (see 5.1.3).

Examples of lo cal IF-divergences

In this section we are going to study which of the examples of IF-dive rgen ces of Section 5.1.3 are in particular lo cal IF-divergences.

Let us b egin with the Hamming distance (see Section 5.1.3). It is defined by:

$$l_{\text{IFS}}(A, B) = \sum_{i=1}^{n} |\mu_A(\omega_i) - \mu_B(\omega_i)| + |\nu_A(\omega_i) - \nu_B(\omega_i)| + |\pi_A(\omega_i) - \pi_B(\omega_i)|.$$

Consider two IF-sets A and B, and an element ω Ω. We have to see that the difference $l_{\text{IFS}}(A, B) - l_{\text{IFS}}(A \; \{\omega\}, B \; \{\omega\})$ only dep ends on $\mu_A(\omega_i), \mu_B(\omega_i), \nu_A(\omega_i)$ and $\nu_B(\omega_i)$. Note that, since $\mu_{A \{\omega_i\}}(\omega_i) = \mu_{B \{\omega_i\}}(\omega_i) = 1$ and $\nu_{A \{\omega_i\}}(\omega_i) = \nu_{B \{\omega_i\}}(\omega_i) = 0$, $l_{\text{IFS}}(A \; \{\omega\}, B \; \{\omega\})$ takes the following value:

$$l_{\text{IFS}}(A \; \{\omega\}, B \; \{\omega\}) = \sum_{j=i} |\mu_A(\omega_j) - \mu_B(\omega_j)| + |\nu_A(\omega_j) - \nu_B(\omega_j)| + |\pi_A(\omega_j) - \pi_B(\omega_j)|$$

whence

$$\begin{aligned} l_{\mathrm{IFS}}(A, B) - l_{\mathrm{IFS}}(A \setminus \{\omega_i\}, B \setminus \{\omega_i\}) = \\ |\mu_A(\omega_i) - \mu_B(\omega_i)| + |\nu_A(\omega_i) - \nu_B(\omega_i)| + |\pi_A(\omega_i) - \pi_B(\omega_i)| = \\ h_{\mathrm{IFS}}(\mu_A(\omega_i), \nu_A(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)). \end{aligned}$$

Thus, l_{IFS} is a local IF-divergence whose associated function h_{IFS} is given by:

$$h_{\mathrm{IFS}}(x_1, x_2, y_1, y_2) = |x_1 - y_1| + |x_2 - y_2| + |x_1 + x_2 - y_1 - y_2|.$$

Moreover, the normalized Hamming distance, defined by $l_{\mathrm{nIFS}}(A, B) = \frac{1}{n} l_{\mathrm{IFS}}(A, B)$, is also a local IF-divergence. The reason is that $l_{\mathrm{nIFS}}(A, B) = \varphi(l_{\mathrm{IFS}}(A, B))$, where $\varphi(x) = \frac{x}{n}$, and we have already mentioned that in that case l_{nIFS} is local if and only if φ is linear.

Let us next study the Hausdorff distance for IF-sets (see Section 5.1.3), which is given by:

$$d_H(A, B) = \sum_{i=1}^{n} \max(|\mu_A(\omega_i) - \mu_B(\omega_i)|, |\nu_A(\omega_i), \nu_B(\omega_i)|).$$

Consider $\omega_i \in \Omega$, and let A and B be two IF-sets. As we have done in the previous case, $d_H(A \setminus \{\omega_i\}, B \setminus \{\omega_i\})$ is given by

$$d_H(A \setminus \{\omega_i\}, B \setminus \{\omega_i\}) = \sum_{j \neq i} \max(|\mu_A(\omega_j) - \mu_B(\omega_j)|, |\nu_A(\omega_j), \nu_B(\omega_j)|),$$

taking into account that $A \setminus \{\omega_i\}$ and $B \setminus \{\omega_i\}$ are given by:

$$\begin{aligned} A \setminus \{\omega_i\} &= \{(\omega_j, \mu_A(\omega_j), \nu_A(\omega_j)), (\omega_i, 1, 0) \mid j \neq i\}. \\ B \setminus \{\omega_i\} &= \{(\omega_j, \mu_B(\omega_j), \nu_B(\omega_j)), (\omega_i, 1, 0) \mid j \neq i\}. \end{aligned}$$

Hence, $d_H(A, B) - d_H(A \setminus \{\omega_i\}, B \setminus \{\omega_i\})$ is given by

$$d_H(A, B) - d_H(A \setminus \{\omega_i\}, B \setminus \{\omega_i\}) = \max(|\mu_A(\omega_i) - \mu_B(\omega_i)|, |\nu_A(\omega_i) - \nu_B(\omega_i)|).$$

Therefore, the Hamming distance for IF-sets is a local IF-divergence, whose associated function h_{d_H} is given by

$$h_{d_H}(x_1, x_2, y_1, y_2) = \max(|x_1 - y_1|, |x_2 - y_2|).$$

The same applies to the normalized Hausdorff distance, since it is a linear transformation of the Hausdorff distance.

Consider now the IF-divergences defined by Hong and Kim, D_C and D_L (see Section 5.1.3), given by

$$D_C(A, B) = \frac{1}{2n} \sum_{i=1}^{n} |\mu_A(\omega_i) - \mu_B(\omega_i)| + |\nu_A(\omega_i) - \nu_B(\omega_i)|.$$

$$\begin{aligned} D_L(A, B) = \frac{1}{4n} \sum_{i=1}^{n} & |\mu_A(\omega_i) - \mu_B(\omega_i) - \nu_A(\omega_i) + \nu_B(\omega_i)| \\ & + |\mu_A(\omega_i) - \mu_B(\omega_i)| + |\nu_A(\omega_i) - \nu_B(\omega_i)|. \end{aligned}$$

Let us see that both IF-divergences are local. Consider two IF-sets A and B and an element $\omega \in \Omega$, and let us compute $D_C(A \cup \{\omega\}, B \cup \{\omega\})$ and $D_L(A \cup \{\omega\}, B \cup \{\omega\})$.

$$D_C(A \cup \{\omega\}, B \cup \{\omega\}) = \frac{1}{2n} \sum_{j \neq i} |\mu_A(\omega_j) - \mu_B(\omega_j)| + |\nu_A(\omega_j) - \nu_B(\omega_j)|.$$

$$D_L(A \cup \{\omega\}, B \cup \{\omega\}) = \frac{1}{4n} \sum_{j \neq i} |\mu_A(\omega_j) - \mu_B(\omega_j) - \nu_A(\omega_j) + \nu_B(\omega_j)| + |\mu_A(\omega_j) - \mu_B(\omega_j)| + |\nu_A(\omega_j) - \nu_B(\omega_j)|.$$

Then,

$$D_C(A, B) - D_C(A \cup \{\omega\}, B \cup \{\omega\}) = |\mu_A(\omega_i) - \mu_B(\omega_i)| + |\nu_A(\omega_i) - \nu_B(\omega_i)|.$$

$$\begin{aligned} D_L(A, B) - D_L(A \cup \{\omega\}, B \cup \{\omega\}) &= |\mu_A(\omega_i) - \mu_B(\omega_i) - \nu_A(\omega_i) + \nu_B(\omega_i)| \\ &+ |\mu_A(\omega_i) - \mu_B(\omega_i)| + |\nu_A(\omega_i) - \nu_B(\omega_i)|. \end{aligned}$$

Thus, both IF-divergences are local, and their respective functions h_{D_C} and h_{D_L} are:

$$h_{D_C}(x_1, x_2, y_1, y_2) = |x_1 - y_1| + |x_2 - y_2|.$$

$$h_{D_L}(x_1, x_2, y_1, y_2) = |x_1 - y_1 - x_2 + y_2| + |x_1 - y_1| + |x_2 - y_2|.$$

In summary, Hamming and Hausdorff distances and the IF-divergences of Hong and Kim are local IF-divergences. It can be checked that the other examples of IF-divergences are not local.

5.1.5 IF-divergences Vs Divergences

Some of the studies presented until now in this chapter are inspired in the concept of fuzzy divergence introduced by Montes et al. ([160]).

Definition 5.40 ([160]) *Let Ω be an universe. A map $D : FS(\Omega) \times FS(\Omega) \to \mathbb{R}$ is a divergence if it satisfies the following conditions:*

Div.1: $D(A, A) = 0$ *for every* $A \in FS(\Omega)$.
Div.2: $D(A, B) = D(B, A)$ *for every* $A, B \in FS(\Omega)$.
Div.3: $D(A \cap C, B \cap C) \leq D(A, B)$*, for every* $A, B, C \in FS(\Omega)$.
Div.4: $D(A \cup C, B \cup C) \leq D(A, B)$*, for every* $A, B, C \in FS(\Omega)$.

Montes et all ([160]) also investigated the local property for fuzzy divergences.

Definition 5.41 ([160, Def. 3.2]) *A divergence measure defined on a finite universe is a local divergence, or it is said to fulfill the local property, if for every $A, B \in FS(\Omega)$ and every $\omega \in \Omega$ we have that:*

$$D(A, B) - D(A \cup \{\omega\}, B \cup \{\omega\}) = h(A(\omega), B(\omega)),$$

Local fuzzy divergences were characterized as follows.

Theorem 5.42 ([160, Prop. 3.4]) *A map $D: FS(\Omega) \times FS(\Omega) \to \mathbb{R}$ defined on a finite universe $\Omega = \{\omega_1, \ldots, \omega_n\}$ is a local divergence if and only if there is a function $h : [0,1] \times [0,1] \to \mathbb{R}$ such that*

$$D(A, B) = \sum_{i=1}^{n} h(A(\omega_i), B(\omega_i)),$$

and

loc.1: $h(x, y) = h(y, x)$, *for every* $(x, y) \in [0, 1]^2$.
loc.2: $h(x, x) = 0$ *for every* $x \in [0, 1]$.
loc.3: $h(x, z) \geq \max(h(x, y), h(y, z))$ *for every* $x, y, z \in [0, 1]$ *such that* $x < y < z$.

In this section we are going to study the relationship between divergences and IF-divergences. We shall provide some methods to derive IF-divergences from divergences and vice versa. Moreover, we shall investigate under which conditions the property of being local is preserved under these transformations.

From IF-divergences to fuzzy divergences

Consider an IF-divergence $D_{IFS} : IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ defined on a finite universe $\Omega = \{\omega_1, \ldots, \omega_n\}$. Recall that every fuzzy set A is in particular an IF-set, whose membership and non-membership functions are $\mu_A(\omega_i) = A(\omega_i)$ and $\nu_A(\omega_i) = 1 - A(\omega_i)$, respectively. Hence, if A and B are two fuzzy sets, we can compute its divergence D as:

$$D(A, B) = D_{IFS}(A, B).$$

Proposition 5.43 *If D_{IFS} is an IF-divergence, the map $D : FS(\Omega) \times FS(\Omega) \to \mathbb{R}$ given by*

$$D(A, B) = D_{IFS}(A, B)$$

is a divergence for fuzzy sets. Moreover, if D_{IFS} satisfies axiom IF-Div.5, then D satisfies axiom Div.5, and if D_{IFS} is local, then so is D.

Proof Let us prove that D is a divergence, i.e., that it satisfies axioms Diss.1 to Div.4.

~~Diss.1~~: Let A be a fuzzy set. Then:

$$D(A, A) = D_{IFS}(A, A) = 0.$$

~~Diss.2:~~ Let A and B be two fuzzy sets. Sincetheyare in particular IF-sets, $D_{\mathrm{IFS}}(A, B) = D_{\mathrm{IFS}}(B, A)$, and therefore:

$$D(A, B) = D_{\mathrm{IFS}}(A, B) = D_{\mathrm{IFS}}(B, A) = D(B, A).$$

~~Div.3:~~ Let A, B and C be fuzzy sets. Again, since theyare inparticular IF-sets, it holds that $D_{\mathrm{IFS}}(A \cap C, B \cap C) \le D_{\mathrm{IFS}}(A, B)$. Then:

$$D(A \cap C, B \cap C) = D_{\mathrm{IFS}}(A \cap C, B \cap C) \le D_{\mathrm{IFS}}(A, B) = D(A, B).$$

~~Div.4:~~Similarly to Div.3, consider fuzzy sets A, B and C. Since they are in particular IF-sets, theysatisfy $D_{\mathrm{IFS}}(A \quad C, B \quad C) \le D_{\mathrm{IFS}}(A, B)$,whence

$$D(A \quad C, B \quad C) = D_{\mathrm{IFS}}(A \quad C, B \quad C) \le D_{\mathrm{IFS}}(A, B) = D(A, B).$$

Thus, D isa divergenceforfuzzy sets. Assume nowthat D_{IFS} satisfiesIF-Div.5, i.e.,

$$D_{\mathrm{IFS}}(A, B) = D_{\mathrm{IFS}}(A^c, B^c) \text{ for every } A, B \quad IFSs(\Omega).$$

Then, in particular, D satisfiesaxiom Div.5

$$D(A, B) = D_{\mathrm{IFS}}(A, B) = D_{\mathrm{IFS}}(A^c, B^c) = D(A^c, B^c),$$

for every $A, B \quad FS(\Omega)$. Assu me now that D_{IFS} is a lo cal IF-divergence.Then:

$$\begin{aligned} D(A, B) - D(A \quad \{\omega\}, B \quad \{\omega\}) &= D_{\mathrm{IFS}}(A, B) - D_{\mathrm{IFS}}(A \quad \{\omega\}, B \quad \{\omega\}) \\ &= h(A(\omega), 1 - A(\omega), B(\omega), 1 - B(\omega)) = h(A(\omega), B(\omega)), \end{aligned}$$

where $h(x, y) = h(x, 1 - x, y, 1 - y)$. Consequently, D is a lo cal divergence between fuzzy sets. ■

Remark 5.44 *The function D definedin the previousproposition isin facta composition of some functions:*

$$D : FS(\Omega) \times FS(\Omega) \xrightarrow{i} IFSs(\Omega) \times IFSs(\Omega) \xrightarrow{D_{\mathrm{IFS}}} \mathbb{R}$$

where $i(A, B)$ stands for the inclusion of $FS(\Omega) \times FS(\Omega)$ on $IFSs(\Omega) \times IFSs(\Omega)$.

Remark 5.45 *IfwelookattheproofofProposition5.43, weseethat, inordertoprove that D satisfiesaxiom Div. i, for $i \quad \{1,2\}$ it is enough for D_{IFS} tosatisfy axiomIF-Diss.i. Moreover, if D_{IFS} satisfies axiomIF-Div. j, for $j \quad \{3,4\}$, then D also satisfies axiom Div.j. In fact, if for instance D_{IFS} is notan IF-divergence, but it satisfiesIF-Diss.1, IF-Diss.2 andIF-Div.3, we cannot assurethat D is a divergence. However, we know that D satisfies axiomsDiv.1, IF-Div.2 andIF-Div.3.*

The above method of deriving divergences from IF-divergences seems to be natural. Let us show how it can be used in a few examples.

Example 5.46 *Consider the Hamming distance for IF-sets that we have already studied in Section 5.1.3, given by:*

$$l_{\mathrm{IFS}}(A,B)=\frac{1}{2}\sum_{i=1}^{n}(|\mu_A(\omega_i)-\mu_B(\omega_i)|+|\nu_A(\omega_i)-\nu_B(\omega_i)|+|\pi_A(\omega_i)-\pi_B(\omega_i)|).$$

If we consider A and B two fuzzy sets, the divergence D defined in the previous proposition is:

$$D_1(A,B)=\sum_{i=1}^{n}|A(\omega_i)-B(\omega_i)|.$$

Recall that the function:

$$l_{\mathrm{FS}}(A,B)=\sum_{i=1}^{n}|A(\omega_i)-B(\omega_i)|,\quad A,B\in FS(\Omega)$$

is known as the Hamming distance for fuzzy sets. Then, from the Hamming distance for IF-sets we obtain the Hamming distance for fuzzy sets. Moreover, if we consider the normalized Hamming distance for IF-sets, we also obtain the normalized Hamming distance, defined by $l_{\mathrm{nFS}}(A,B)=\frac{1}{n}l_{\mathrm{FS}}$, for fuzzy sets.

Consider now the Hausdorff distance (see Section 5.1.3) for IF-sets:

$$d_{\mathrm{H}}(A,B)=\sum_{i=1}^{n}\max(|\mu_A(\omega_i)-\mu_B(\omega_i)|,|\nu_A(\omega_i)-\nu_B(\omega_i)|).$$

Given two fuzzy sets A and B, if we apply Proposition 5.43 we obtain the Hamming distance for fuzzy sets:

$$\begin{aligned}D_2(A,B)=d_{\mathrm{H}}(A,B)&=\sum_{i=1}^{n}\max(|A(\omega_i)-B(\omega_i)|,|(1-A(\omega_i))-(1-B(\omega_i))|)\\&=\sum_{i=1}^{n}|A(\omega_i)-B(\omega_i)|=l_{\mathrm{FS}}(A,B).\end{aligned}$$

Moreover, if we consider the normalized Hausdorff distance, we obtain the normalized Hamming distance:

$$\begin{aligned}D_3(A,B)=d_{\mathrm{nH}}(A,B)&=\frac{1}{n}\sum_{i=1}^{n}\max(|A(\omega_i)-B(\omega_i)|,|(1-A(\omega_i))-(1-B(\omega_i))|)\\&=\frac{1}{n}\sum_{i=1}^{n}|A(\omega_i)-B(\omega_i)|=l_{\mathrm{nFS}}(A,B).\end{aligned}$$

Thus, both the Hamming distance and the Hausdorff distance for IF-sets produce the same divergence for fuzzy sets: the Hamming distance for fuzzy sets.

However, if we consider the IF-divergences of Hong and Kim (see Section 5.1.3), defined by:

$$D_C(A,B) = \frac{1}{2n}\sum_{i=1}^{n}\left(|\mu_A(\omega_i) - \mu_B(\omega_i)| + |\nu_A(\omega_i) - \nu_B(\omega_i)|\right);$$

$$D_L(A,B) = \frac{1}{4n}\sum_{i=1}^{n}\Big(|(\mu_A(\omega_i) - \mu_B(\omega_i)) - (\nu_A(\omega_i) - \nu_B(\omega_i))| + |\mu_A(\omega_i) - \mu_B(\omega_i)| + |\nu_A(\omega_i) - \nu_B(\omega_i)|\Big);$$

and we apply Proposition 5.43 we obtain also the normalized Hamming distance:

$$\begin{aligned}
D_4(A,B) = D_C(A,B) &= \frac{1}{2n}\sum_{i=1}^{n}\left(|A(\omega_i) - B(\omega_i)| + |(1 - A(\omega_i)) - (1 - B(\omega_i))|\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}|A(\omega_i) - B(\omega_i)| = l_{nFS}(A,B). \\
D_5(A,B) = D_L(A,B) &= \frac{1}{4n}\sum_{i=1}^{n}\Big(|(A(\omega_i) - B(\omega_i)) - (1 - A(\omega_i) - 1 + B(\omega_i))| \\
&\quad + |A(\omega_i) - B(\omega_i)| + |1 - A(\omega_i) - 1 + B(\omega_i)|\Big) \\
&= \frac{1}{n}\sum_{i=1}^{n}|A(\omega_i) - B(\omega_i)| = l_{nFS}(A,B).
\end{aligned}$$

Thus, both Hamming and Hausdorff distances for IF-sets produce the Hamming distance for fuzzy sets, and the normalized Hamming and Hausdorff distances, and Hong and Kim dissimilarities for IF-sets produce the normalized Hamming distance for fuzzy sets. Consequently, all these IF-divergences can be seen as generalizations of the Hamming distance for fuzzy sets to the comparison of IF-sets.

Example 5.47 *Let us now consider the IF-divergence defined by Li et al. (see page 283 of Section 5.1.3):*

$$D_O(A,B) = \sqrt{\frac{1}{2n}}\left(\sum_{i=1}^{n}(\mu_A(\omega_i) - \mu_B(\omega_i))^2 + (\nu_A(\omega_i) - \nu_B(\omega_i))^2\right)^{\frac{1}{2}}.$$

If we use Proposition 5.43 in order to build a divergence for fuzzy sets from D_O, we obtain the normalized Euclidean distance for fuzzy sets:

$$\begin{aligned}
D(A,B) = D_O(A,B) &= \sqrt{\frac{1}{2n}}\left(\sum_{i=1}^{n}(A(\omega_i) - B(\omega_i))^2 + (1 - A(\omega_i) - 1 + B(\omega_i))^2\right)^{\frac{1}{2}} \\
&= \sqrt{\frac{1}{n}}\left(\sum_{i=1}^{n}(A(\omega_i) - B(\omega_i))^2\right)^{\frac{1}{2}} = \sqrt{2}\,q_{nFS}(A,B).
\end{aligned}$$

Thus, both the normalized Euclidean distance for IF-sets and Li et al. IF-divergence are generalizations of the normalized Euclidean distance for fuzzy sets. Note however that the normalized Euclidean distance is not an IF-divergence (see Section 5.1.3), even though Li et al.'s dissimilarity is.

Example 5.48 *Consider now the IF-divergence defined by Mitchell (see Section 5.1.3):*

$$D_{\mathrm{HB}}(A,B) = \frac{1}{2\sqrt[p]{n}}\left[\left(\sum_{i=1}^{n}|\mu_{\mathrm{A}}(\omega_i)-\mu_{\mathrm{B}}(\omega_i)|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n}|\nu_{\mathrm{A}}(\omega_i)-\nu_{\mathrm{B}}(\omega_i)|^p\right)^{\frac{1}{p}}\right].$$

Applying Proposition 5.43, we obtain the following divergence for fuzzy sets:

$$\begin{aligned} D_1(A,B) &= D_{\mathrm{HB}}(A,B) = \frac{1}{2\sqrt[p]{n}}\left[\left(\sum_{i=1}^{n}|A(\omega_i)-B(\omega_i)|^p\right)^{\frac{1}{p}} \right.\\ &\left. + \left(\sum_{i=1}^{n}|(1-A(\omega_i))-(1-B(\omega_i))|^p\right)^{\frac{1}{p}}\right] = \frac{1}{\sqrt[p]{n}}\left(\sum_{i=1}^{n}|A(\omega_i)-B(\omega_i)|^p\right)^{\frac{1}{p}}. \end{aligned}$$

If we now consider the IF-Divergence D_{e}^p of Liang and Shi (see Section 5.1.3), defined by:

$$D_{\mathrm{e}}^p(A,B) = \frac{1}{2\sqrt[p]{n}}\left(\sum_{i=1}^{n}\left(|\mu_{\mathrm{A}}(\omega_i)-\mu_{\mathrm{B}}(\omega_i)| + |\nu_{\mathrm{A}}(\omega_i)-\nu_{\mathrm{B}}(\omega_i)|\right)^p\right)^{\frac{1}{p}}$$

and apply Proposition 5.43, we obtain the following divergence:

$$\begin{aligned} D_2(A,B) &= D_{\mathrm{e}}^p(A,B) \\ &= \frac{1}{2\sqrt[p]{n}}\left(\sum_{i=1}^{n}\left(|A(\omega_i)-B(\omega_i)| + |(1-A(\omega_i))-(1-B(\omega_i))|\right)^p\right)^{\frac{1}{p}} \\ &= \frac{1}{\sqrt[p]{n}}\left(\sum_{i=1}^{n}|A(\omega_i)-B(\omega_i)|^p\right)^{\frac{1}{p}}. \end{aligned}$$

Note that $D_1(A,B) = D_2(A,B)$. Thus, both D_{HB} and D_{e}^p produce the same divergence between fuzzy sets, and therefore both of them can be seen as a generalization of the divergence D_1.

Although the method proposed in Proposition 5.43 seems to be very natural, there is another possible, albeit less intuitive, way of deriving divergences from IF-divergences, that we detail next.

Proposition 5.49 *The function $D : FS(\Omega) \times FS(\Omega) \to \mathbb{R}$ defined by*

$$D(A,B) = D_{\mathrm{IFS}}(A,B),$$

where D_{IFS} *isan IF-divergence,is a divergencefor fuzzysets, where* A *and* B *are given by:*

$$A = \{(\omega, A(\omega), 0) | \omega \quad \Omega\} \quad IFSs(\Omega).$$
$$B = \{(\omega, B(\omega), 0) | \omega \quad \Omega\} \quad IFSs(\Omega).$$

However, although D_{IFS} *satisfies IF-Div.5,* D *may notsatisfy Div.5.*

Pro of Letussee that D satisfiesthe divergenceaxioms.

~~Diss.1~~: Let A b e a fuzzy se t.Then $A = \{(\omega, A(\omega), 0) | \omega \quad \Omega\}$, and therefore, as D_{IFS} isan IF-divergence,

$$D(A,A) = D_{\text{IFS}}(A, A) = 0.$$

~~Diss.2~~: Let A and B be two fuzzy sets. Then

$$D(A, B) = D_{\text{IFS}}(A, B) = D_{\text{IFS}}(B, A) = D(B, A),$$

because D_{IFS} is symmetric.

~~Div.3:~~ Consider $A, B, C \quad IFSs(\Omega)$. Since D_{IFS} is anIF-divergence, $D_{\text{IFS}}(A \cap C, B \cap C) \leq D_{\text{IFS}}(A, B)$. Moreover,

$$A \cap C = \{(\omega, \min(\mu_A(\omega), \mu_C(\omega)), 0) | \omega \quad \Omega\} = A \cap C.$$
$$B \cap C = \{(\omega, \min(\mu_B(\omega), \mu_C(\omega)), 0) | \omega \quad \Omega\} = B \cap C,$$

whence

$$\begin{aligned} D(A \cap C, B \cap C) &= D_{\text{IFS}}(A \cap C, B \cap C) \\ &= D_{\text{IFS}}(A \cap C, B \cap C) \leq D_{\text{IFS}}(A, B) = D(A, B). \end{aligned}$$

~~Div.4:~~ The pro of is similar to the previous one. Consider three fuzzy sets A, B and C. We know that $D_{\text{IFS}}(A \quad C, B \quad C) \leq D_{\text{IFS}}(A, B)$. Moreover,

$$A \quad C = \{(\omega, \max(\mu_A(\omega), \mu_C(\omega)), 0) | \omega \quad \Omega\} = A \quad C.$$
$$B \quad C = \{(\omega, \max(\mu_B(\omega), \mu_C(\omega)), 0) | \omega \quad \Omega\} = B \quad C.$$

Then, axiom Div.4 is satisfied, because:

$$\begin{aligned} D(A \quad C, B \quad C) &= D_{\text{IFS}}(A \quad C, B \quad C) \\ &= D_{\text{IFS}}(A \quad C, B \quad C) \leq D_{\text{IFS}}(A, B) = D(A, B). \end{aligned}$$

Hence, D is a divergence for fu zzy setsAs sume now that D_{IFS} satisfiesaxiom IF-Div.5 and let us show thatin that case D may notsatisfy Div.5. Considera singleton universe $\Omega = \{\omega\}$,and the function $D_{\text{IFS}} : IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ defined by

$$\begin{aligned} D_{\text{IFS}}(A, B) = &\left| \max(\mu_A(\omega) - 0.5, 0) - \max(\mu_B(\omega) - 0.5, 0) \right| \\ + &\left| \max(\nu_A(\omega) - 0.5, 0) - \max(\nu_B(\omega) - 0.5, 0) \right| \end{aligned}$$

Let us see that D_{IFS} is an IF-divergence.

~~IF-Diss.1:~~ Let A be an IF-set. Trivially

$$|\max(\mu_A(\omega) - 0.5, 0) - \max(\mu_A(\omega) - 0.5, 0)| = 0 \quad \text{and}$$
$$|\max(\nu_A(\omega) - 0.5, 0) - \max(\nu_A(\omega) - 0.5, 0)| = 0,$$

and therefore $D_{IFS}(A, A) = 0$.

~~IF-Diss.2:~~ Let A and B be two IF-sets. Then it follows from the definition that $D_{IFS}(A, B) = D_{IFS}(B, A)$.

~~IF-Div.3:~~ Let A, B and C be three IF-sets. We must prove the following inequality:

$$\begin{aligned}
&|\max(\mu_A(\omega) - 0.5, 0) - \max(\mu_B(\omega) - 0.5, 0)| + \\
&|\max(\nu_A(\omega) - 0.5, 0) - \max(\nu_B(\omega) - 0.5, 0)| \geq \\
&\quad |\max(\mu_{A\cap C}(\omega) - 0.5, 0) - \max(\mu_{B\cap C}(\omega) - 0.5, 0)| + \\
&\quad |\max(\nu_{A\cap C}(\omega) - 0.5, 0) - \max(\nu_{B\cap C}(\omega) - 0.5, 0)|.
\end{aligned}$$

This follows from Lemma A.5 in Appendix A.

~~IF-Div.4:~~ Similarly, if A, B and C are three IF-sets, condition IF-Div.4 holds if and only if:

$$\begin{aligned}
&|\max(\mu_A(\omega) - 0.5, 0) - \max(\mu_B(\omega) - 0.5, 0)| + \\
&|\max(\nu_A(\omega) - 0.5, 0) - \max(\nu_B(\omega) - 0.5, 0)| \geq \\
&\quad |\max(\mu_{A\;C}(\omega) - 0.5, 0) - \max(\mu_{B\;C}(\omega) - 0.5, 0)| + \\
&\quad |\max(\nu_{A\;C}(\omega) - 0.5, 0) - \max(\nu_{B\;C}(\omega) - 0.5, 0)|,
\end{aligned}$$

and this follows from Lemma A.5 in Appendix A.

Hence, D_{IFS} is an IF-divergence. Moreover, it also trivially satisfies axiom IF-Div.5.

Consider the divergence derived in this proposition:

$$D(A, B) = D_{IFS}(\{(\omega, A(\omega), 0)\}, \{(\omega, B(\omega), 0)\}) = |\max(A(\omega) - 0.5, 0) - \max(B(\omega) - 0.5, 0)|$$

Although D_{IFS} satisfies IF-Div.5, D does not fulfill Div.5: if we consider the fuzzy sets A and B given by

$$\begin{aligned}
A &= \{(\omega, 0.3)\} & A^c &= \{(\omega, 0.7)\}, \text{ and} \\
B &= \{(\omega, 0.4)\} & B^c &= \{(\omega, 0.6)\}
\end{aligned}$$

then it holds that $D(A, B) = 0 = 0.1 = D(A^c, B^c)$. ∎

Although this second method for deriving divergences from IF-divergences is also valid, for us the first one seems to be more natural; besides, we have shown that some of the most important examples of divergences can be obtained applying this method to the corresponding IF-divergences.

From fuzzy divergences to IF-divergences

Consider now a divergence $D: FS(\Omega) \times FS(\Omega) \to \mathbb{R}$ between fuzzy sets define on a finite space $\Omega = \{\omega_1, \ldots, \omega_n\}$, and let us study how to derive an IF-divergence from it. Consider two IF-sets A and B. Each of them can be decomposed into two fuzzy sets as follows:

$$A = \{(\omega_i, \mu_A(\omega_i), \nu_A(\omega_i)) \mid i = 1, \ldots, n\} \quad IFSs(\Omega)$$
$$A_1 = \{(\omega_i, \mu_A(\omega_i)) \mid i = 1, \ldots, n\} \quad FS(\Omega) \quad IFSs(\Omega).$$
$$A_2 = \{(\omega_i, \nu_A(\omega_i)) \mid i = 1, \ldots, n\} \quad FS(\Omega) \quad IFSs(\Omega).$$
$$B = \{(\omega_i, \mu_B(\omega_i), \nu_B(\omega_i)) \mid i = 1, \ldots, n\} \quad IFSs(\Omega)$$
$$B_1 = \{(\omega_i, \mu_B(\omega_i)) \mid i = 1, \ldots, n\} \quad FS(\Omega) \quad IFSs(\Omega).$$
$$B_2 = \{(\omega_i, \nu_B(\omega_i)) \mid i = 1, \ldots, n\} \quad FS(\Omega) \quad IFSs(\Omega).$$

Using the divergence D we can measure the divergence between the pairs of fuzzy sets (A_1, B_1) and (A_2, B_2). In other words, we have the divergence between the membership degrees and the non-membership degrees; in order to compute the divergence between A and B it only remains to combine these two divergences.

Theorem 5.50 *Let D be a divergence for fuzzy sets, and let $f : [0, \infty) \times [0, \infty) \to [0, \infty)$ be a mapping satisfying the following two properties:*

- *f1:* $f(0,0) = 0$;
- *f2:* $f(\ ,t)$ *and* $f(t,\ )$ *are increasing for every* $t \ [0, \infty)$;

then, the function $D_{IFS} : IFSs(\Omega) \times IFSs(\Omega) \to \mathbb{R}$ *defined by*

$$D_{IFS}(A, B) = f(D(A_1, B_1), D(A_2, B_2)), \quad \text{for every } A, B \quad IFSs(\Omega),$$

is an IF-divergence. Moreover, if D is a local divergence, then D_{IFS} is also a local IF-divergence iff f has the form: $f(x, y) = \alpha x + \beta y$, for some $\alpha, \beta \geq 0$.

Finally, if f is symmetric then D_{IFS} fulfills axiom IF-Div.5 (regardless of whether D satisfies or not axiom Div.5), and if f is not symmetric, then although D satisfies Div.5, D_{IFS} may not satisfy IF-Div.5.

Pro of We begin by showing that D_{IFS} is an IF-divergence.

~~IF-Diss.1:~~ Let A be an IF-set. Applying the definition of D_{IFS} we obtain that:

$$D_{IFS}(A, A) = f(D(A_1, A_1), D(A_2, A_2)) = f(0, 0) \overset{f1}{=} 0.$$

~~IF-Diss.2:~~ Let A, B be IF-sets, and let us prove that $D_{IFS}(A, B) = D_{IFS}(B, A)$.

$$\begin{aligned} D_{IFS}(A, B) &= f(D(A_1, B_1), D(A_2, B_2)) \\ &= f(D(B_1, A_1), D(B_2, A_2)) = D_{IFS}(B, A). \end{aligned}$$

IF-Div.3: Consider the IF-sets A, B and C, and let us prove that $D_{IFS}(A \cap C, B \cap C) \leq D_{IFS}(A, B)$. Let us note the following:

$$\begin{aligned} A \cap C &= \{(\omega, \mu_{A \cap C}(\omega), \nu_{A \cap C}(\omega)) \mid \omega \in \Omega\} \\ &= \{(\omega, \min(\mu_A(\omega), \mu_C(\omega)), \max(\nu_A(\omega), \nu_C(\omega))) \mid \omega \in \Omega\} \\ (A \cap C)_1 &= \{(\omega, \min(\mu_A(\omega), \mu_C(\omega))) \mid \omega \in \Omega\} \subseteq FS(\Omega). \\ (A \cap C)_2 &= \{(\omega, \max(\nu_A(\omega), \nu_C(\omega))) \mid \omega \in \Omega\} \subseteq FS(\Omega). \end{aligned}$$

Similarly, we also obtain that

$$\begin{aligned} (B \cap C)_1 &= \{(\omega, \min(\mu_B(\omega), \mu_C(\omega))) \mid \omega \in \Omega\} \subseteq FS(\Omega). \\ (B \cap C)_2 &= \{(\omega, \max(\nu_B(\omega), \nu_C(\omega))) \mid \omega \in \Omega\} \subseteq FS(\Omega). \end{aligned}$$

Since D is a divergence for fuzzy sets, applying Div.3 we obtain that:

$$D(A \cap C_1, B \cap C_1) = D((A \cap C)_1, (B \cap C)_1) \leq D(A_1, B_1),$$

where $C_1 = \mu_C$, and applying Div.4,

$$D(A \cup C_2, B \cup C_2) = D((A \cap C)_2, (B \cap C)_2) \leq D(A_2, B_2),$$

where $C_2 = \nu_C$. From these properties, $D_{IFS}(A \cap C, B \cap C) \leq D_{IFS}(A, B)$

$$\begin{aligned} D_{IFS}(A \cap C, B \cap C) &= f(D((A \cap C)_1, (B \cap C)_1), D((A \cup C)_2, (B \cup C)_2)) \\ &\leq f(D(A_1, B_1), D(A_2, B_2)) = D_{IFS}(A, B). \end{aligned}$$

IF-Div.4: Let us prove that $D_{IFS}(A \cup C, B \cup C) \leq D_{IFS}(A, B)$ for every IF-sets A, B and C, similarly to the previous point. We have that

$$\begin{aligned} A \cup C_1 &= \{(\omega, \max(\mu_A(\omega), \mu_C(\omega))) \mid \omega \in \Omega\} \subseteq FS(\Omega), \\ A \cup C_2 &= \{(\omega, \min(\nu_A(\omega), \nu_C(\omega))) \mid \omega \in \Omega\} \subseteq FS(\Omega), \\ B \cup C_1 &= \{(\omega, \max(\mu_B(\omega), \mu_C(\omega))) \mid \omega \in \Omega\} \subseteq FS(\Omega), \\ B \cup C_2 &= \{(\omega, \min(\nu_B(\omega), \nu_C(\omega))) \mid \omega \in \Omega\} \subseteq FS(\Omega). \end{aligned}$$

Applying Div.4,

$$D(A \cup C_1, B \cup C_1) \leq D(A_1, B_1),$$

and Div.3 implies that:

$$D(A \cup C_2, B \cup C_2) \leq D(A_2, B_2).$$

Using these two inequalities, we can prove that $D_{IFS}(A \cup C, B \cup C) \leq D_{IFS}(A, B)$

$$\begin{aligned} D_{IFS}(A \cup C, B \cup C) &= f(D(A \cup C_1, B \cup C_1), D(A \cup C_2, B \cup C_2)) \\ &\leq f(D(A_1, B_1), D(A_2, B_2)) = D_{IFS}(A, B). \end{aligned}$$

Hence, D_{IFS} is an IF-divergence. Assume now that f is symmetric, i.e., $f(x, y) = f(y, x)$ for every $(x, y) \in [0, 1]^2$, then it is immediate that D_{IFS} satisfies axiom IF-Div.5, that is, $D_{\text{IFS}}(A, B) = D_{\text{IFS}}(A^c, B^c)$ for every $A, B \in IFSs(\Omega)$, since:

$$D_{\text{IFS}}(A, B) = f(D(A_1, B_1), D(A_2, B_2)) = f(D(A_2, B_2), D(A_1, B_1)) = D_{\text{IFS}}(A^c, B^c).$$

However, assume that f is not symmetric, and let us give an example of divergence D that fulfills axiom Div.5, such that D_{IFS} does not satisfies IF-Div.5. Consider the normalized Hamming divergence for fuzzy sets:

$$l_{\text{FS}}(A, B) = \frac{1}{n}\sum_{i=1}^{n} |A(\omega_i) - B(\omega_i)|,$$

and let f be given by: $f(x, y) = \alpha x + \beta y$, where $\alpha \neq \beta$, for example $\alpha = 1$ and $\beta = 0$. Then:

$$D_{\text{IFS}}(A, B) = \frac{1}{n}\sum_{i=1}^{n} (\alpha|\mu_A(\omega_i) - \mu_B(\omega_i)| + \beta\,|\nu_A(\omega_i) - \nu_B(\omega_i)|)$$

is an IF-divergence. Obviously D satisfies axiom Div.5, but D_{IFS} does not satisfy IF-Div.5; to see this, it suffices to consider the IF-sets

$$A = \{(\omega, 0.6, 0.2) \mid \omega \in \Omega\} \text{ and } B = \{(\omega, 0.5, 0.4) \mid \omega \in \Omega\}.$$

Then it holds that

$$D_{\text{IFS}}(A, B) = \frac{1}{n}\sum_{i=1}^{n} (\alpha \cdot 0.1 + \beta \cdot 0.2) = \alpha \cdot 0.1 + \beta \cdot 0.2 = 0.1.$$

$$D_{\text{IFS}}(A^c, B^c) = \frac{1}{n}\sum_{i=1}^{n} (\alpha \cdot 0.2 + \beta \cdot 0.1) = \alpha \cdot 0.2 + \beta \cdot 0.1 = 0.2.$$

and therefore $D_{\text{IFS}}(A, B) \neq D_{\text{IFS}}(A^c, B^c)$.

Assume now that D is a local divergence, i.e., that there is a function h, such that

loc.1: $h(x, y) = h(y, x)$, for every $(x, y) \in [0, 1]^2$;
loc.2: $h(x, x) = 0$ for every $x \in [0, 1]$;
loc.3: $h(x, z) \geq \max(h(x, y), h(y, z))$, for every $x, y, z \in [0, 1]$ such that $x < y < z$;

for which D can be expressed by:

$$D(A, B) = \sum_{i=1}^{n} h(A(\omega_i), B(\omega_i)).$$

Then, D_{IFS} is given by

$$D_{\text{IFS}}(A, B) = f\left(\sum_{i=1}^{n} h(\mu_A(\omega_i), \mu_B(\omega_i)), \sum_{i=1}^{n} h(\nu_A(\omega_i), \nu_B(\omega_i))\right).$$

Let us see that if f is linearthen D_{IFS} is a lo cal IF-divergence.In such a case,D_{IFS} has the followi ng form:

$$D_{\mathrm{IFS}}(A, B) = \sum_{i=1}^{n} \alpha \quad h(\mu_A(\omega_i), \mu_B(\omega_i)) + \beta \quad h(\nu_A(\omega_i), \nu_B(\omega_i)),$$

and if we define h by:

$$h(x_1, y_1, x_2, y_2) = \alpha \quad h(x_1, x_2) + \beta \quad h(y_1, y_2)$$

then if suffices to showthat h satisfies prop erties (i)-(iv) to deduce that D_{IFS} is a lo cal IF-divergence. Letussee thatthis indeedthecase:

~~IF-lo c.1~~: Consider$(x, y) \quad [0, 1]^2$. By hyp othesis it holds that $h(x, x) = h(y, y) = 0$, and then:

$$h(x, y, x, y) = \alpha h(x, x) + \beta h(y, y) = 0.$$

~~IF-lo c.2~~: Consider (x_1, x_2) and (y_1, y_2) in T. Then $h(x_1, y_1) = h(y_1, x_1)$ and $h(x_2, y_2) = h(y_2, x_2)$, whence

$$\begin{aligned} h(x_1, x_2, y_1, y_2) &= \alpha h(x_1, y_1) + \beta h(x_2, y_2) \\ &= \alpha h(y_1, x_1) + \beta h(y_2, x_2) = h(y_1, y_2, x_1, x_2). \end{aligned}$$

~~IF-lo c.3~~: Take now$(x_1, x_2), (y_1, y_2) \quad T$ and $z \quad [0, 1]$such that $x_1 \leq z \leq y_1$. Then, lo c.3 implies that:

$$h(x_1, y_1) \geq \max(h(x_1, z), h(z, y_1)),$$

whence

$$\begin{aligned} h(x_1, x_2, y_1, y_2) = \alpha \quad h(x_1, y_1) + \beta \quad h(x_2, y_2) &\geq \alpha \quad \max(h(x_1, z), h(z, y_1)) + \beta \quad h(x_2, y_2) \\ &= \max(h(x_1, x_2, z, y_2), h(z, x_2, y_1, y_2)). \end{aligned}$$

In particular, $h(x_1, x_2, y_1, y_2) \geq h(x_1, x_2, z, y_2)$ and, if $(x_2, z), (y_2, z) \quad T$, then $\max(x_2 + z, y_2 + z) \leq 1$ and $h(x_1, x_2, y_1, y_2) \geq h(z, x_2, y_1, y_2)$.

~~IF-lo c.4~~: Consider $(x_1, x_2), (y_1, y_2) \quad T$ and $z \quad [0, 1]$ such that $x_2 \leq z \leq y_2$. Applying prop erty lo c.3 we see that

$$h(x_2, y_2) \geq \max(h(x_2, z), h(z, y_2))$$

and therefore

$$\begin{aligned} h(x_1, x_2, y_1, y_2) = \alpha h(x_1, y_1) + \beta h(x_2, y_2) &\geq \alpha h(x_1, y_1) + \beta \max(h(x_2, z), h(z, y_2)) \\ &= \max(h(x_1, x_2, y_1, z), h(x_1, z, y_1, y_2)). \end{aligned}$$

~~IF-lo c.5~~: Consider (x_1, x_2), (y_1, y_2) T and z $[0, 1]$ By lo c.1, we knowthat $h(z, z) = 0$. Then:

$$\begin{aligned} h(z, x_2, z, y_2) &= \alpha h(z, z) + \beta h(x_2, y_2) = \beta h(x_2, y_2) \\ &\leq \alpha h(x_1, y_1) + \beta h(x_2, y_2) = h(x_1, x_2, y_1, y_2). \\ h(x_1, z, y_1, z) &= \alpha h(x_1, y_1) + \beta h(z, z) = \alpha h(x_1, y_1) \\ &\leq \alpha h(x_1, y_1) + \beta h(x_2, y_2) = h(x_1, x_2, y_1, y_2). \end{aligned}$$

Thus, D_{IFS} is a lo cal divergence. ∎

Remark 5.51 *In a similar way, it ispossible to prove that, if D_1 and D_2 are two divergences for fuzzy sets,and if $f : [0, \infty) \times [0, \infty) \to [0, \infty)$ is an increasing function with $f(0, 0) = 0$, then the function $D_{\mathrm{IFS}} : IF\,Ss(\Omega) \times IF\,Ss(\Omega) \to \mathbb{R}$ defined by:*

$$D_{\mathrm{IFS}}(A, B) = f(D_1(\mu_A, \mu_B), D_2(\nu_A, \nu_B))$$

for every A, B $IF\,Ss(\Omega)$, is an IF-divergence.

If in particular we conside r the function $f(x, y) = x$ we obtain the following result.

Corollary 5.52 *Let D be amap $D : F\,S(\Omega) \times F\,S(\Omega) \to \mathbb{R}$, and consider thefunction $f : [0, \infty) \times [0, \infty) \to [0, \infty)$ given by $f(x, y) = x$. Define $D_{\mathrm{IFS}} : IF\,Ss(\Omega) \times IF\,S\,s(\Omega) \to \mathbb{R}$ by:*

$$D_{\mathrm{IFS}}(A, B) = f(D(A_1, B_1), D(A_2, B_2)), \quad \text{for every } A, B \quad I\,F\,S\,s(\Omega).$$

Then, if D satisfies axiom Div.i $(i \{1, 2\})$, then D_{IFS} satisfies axiom IF-Diss.i ,and if D satisfies axiomDiv. j $(j \{3, 4\})$, D_{IFS} satisfies axiom IF-v.j . Inparticular, if D is a divergence for fuzzysets, then D_{IFS} is anIF-divergence. Moreover, if D islocal, then D_{IFS} isalso a local IF-divergence. However, D_{IFS} may not satisfy the property IF-D iv.5 even if D satisfies Div.5.

Pro of

- Letus assumethat D satisfies Dis s.1.Then, D_{IFS} satisfies IF-Diss .1 since:

$$D_{\mathrm{IFS}}(A, A) = D(A_1, A_1) = 0.$$

- Letus assumethat D satisfies Dis s.2.Then, D_{IFS} isalso symmetricsince:

$$D_{\mathrm{IFS}}(A, B) = D(A_1, B_1) = D(B_1, A_1) = D_{\mathrm{IFS}}(B, A).$$

- Let us assu me that D satisfies Div.3, and letus see that $D_{\mathrm{IFS}}(A \cap C, B \cap C) \leq D_{\mathrm{IFS}}(A, B)$ for every IF-sets A, B and C.

$$D_{\mathrm{IFS}}(A \cap C, B \cap C) = D(A \cap C_1, B \cap C_1) \leq D(A_1, B_1) = D_{\mathrm{IFS}}(A, B).$$

- Finally, assume that D satisfies Div.4. Then also D_{IFS} satisfies axiom IF-Div.4, since for every A, B and C itholds that:

$$D_{\text{IFS}}(A \cap C, B \cap C) = D(A_1 \cap C_1, B_1 \cap C_1) \leq D(A_1, B_1) = D_{\text{IFS}}(A, B).$$

Thus, if D is adivergence for fuzzy sets, then D_{IFS} is also an IF -divergence.Moreover, taking into account theprevious theoremandthat f isa linearfunction, if D is a lo cal divergence,then D_{IFS} is also a lo cal IF-divergence. Furthermore, wehaveseen inthat result that asufficient condition for D_{IFS} tosatisfy IF-Div.5 isthat f issymmetric, which is not the case for $f(x, y) = x$. Then, wecannot assure D_{IFS} tosatisfy IF-Div.5. ■

Using the previous resultswecan give some examples of IF-divergences.

Example 5.53 *Consider thefunction* $D : FS(\Omega) \times FS(\Omega) \to \mathbb{R}$ *defined by:*

$$D(A, B) = \sum_{\omega \in \Omega} h(A(\omega), B(\omega)),$$

where $h : \mathbb{R}^2 \to \mathbb{R}$ *is given by:*

$$h(x, y) = \begin{cases} 0 & \text{if } x = y. \\ 1 - xy & \text{if } x \neq y. \end{cases}$$

Montes proved that thisfunctionsatisfies Div.1, Div.2 and Div.3(see[159]). Then, if we apply Theorem 5.50with the function $f(x, y) = x$ *, we conclude that the function* D_1 *satisfies IF-Div.1, IF-Div.2 and IF-Div.3.*

Similarly, we can consider the function

$$h(x, y) = \begin{cases} 0 & \text{if } x = y, \\ xy & \text{if } x \neq y, \end{cases}$$

and $D : FS(\Omega) \times FS(\Omega) \to \mathbb{R}$ *defined by:*

$$D(A, B) = \sum_{\omega \in \Omega} h(A(\omega), B(\omega)).$$

Montes et al. ([159]) provedthat D *satisfies Div.1, Div.2and Div.4. Then, applying Theorem 5.50 with the funct ion* $f(x, y) = x$ *, we conclude that the function* D_2 *they generate sat isfies IF-Diss.1, IF-Diss.2 and IF-Div.4.*

These two functions D_1 *and* D_2 *were usedin Example 5.22, and there we have proved that they are not IF-divergences.*

Example 5.54 *In Equation (5.2), we considered a function $D : FS(\Omega) \times FS(\Omega) \to \mathbb{R}$ defined on the space $\Omega = \{\omega\}$ by:*

$$D(A, B) = D_{IFS}(A, B) = |\mu_A(\omega) - \mu_B(\omega)| + |\nu_A(\omega) - \nu_B(\omega)|^2.$$

The Hamming distance for fuzzy sets, l_{FS}, is known to be a divergence for fuzzy sets. Then, applying Theorem 5.50 to this divergence and the function $f(x, y) = x + y^2$, we obtain the function of Equation (5.2), and therefore we conclude that it is an IF-divergence.

Assume now that we have an IF-divergence D_{IFS}. Using Theorem 5.50 we can build a divergence D for fuzzy sets. On the other hand, Proposition 5.43 allows us to derive another IF-divergence D_{IFS}. We next investigate under which conditions these two IF-divergences coincide.

Remark 5.55 *Let us consider D_{IFS} an IF-divergence. Let D be the divergence determined by Proposition 5.43:*

$$D(A, B) = D_{IFS}(A, B), \text{ for every } A, B \quad FS(\Omega).$$

and let D_{IFS} be the IF-divergence derived from D by means of Theorem 5.50:

$$D_{IFS}(A, B) = f(D(A_1, B_1), D(A_2, B_2)), \text{ for every } A, B \quad IFSs(\Omega).$$

Then, $D_{IFS} = D_{IFS}$ if and only if for every $A, B \quad IFSs(\Omega)$ it holds that:

$$D_{IFS}(A, B) = f(D_{IFS}(A_1, B_1), D_{IFS}(A_2, B_2)).$$

Similarly, let D be a divergence for fuzzy sets. Using Theorem 5.50 we can build an IF-divergence D_{IFS}, and applying Proposition 5.43, from D_{IFS} we can derive a divergence D. Again, we want to determine if we recover our initial divergence.

Theorem 5.56 *Let D be a divergence for fuzzy sets, and let D_{IFS} be the IF-divergence derived from D by means of Theorem 5.50, given by*

$$D_{IFS}(A, B) = f(D(A_1, B_1), D(A_2, B_2)) \quad A, B \quad IFSs(\Omega).$$

Let D be the divergence derived from D_{IFS} by means of Proposition 5.43:

$$D(A, B) = D_{IFS}(A, B), \text{ for every } A, B \quad FS(\Omega).$$

Then, $D = D$ if and only if $f(x, y) = x$ for every $(x, y) \quad [0, 1]^2$.

Pro of Letuscompute theexpressionof D :

$$D\ (A, B) = D_{\mathrm{IFS}}(A, B) = f\,(D\,(A, B), D\,(A^c, B^c))),$$

for every A,B $FS(\Omega)$. Thus, $D\,(A, B) = D\ (A, B)$ for every A,B $FS(\Omega)$ if and only if:

$$D(A, B) = f\,(D(A,B), D(A^c, B^c)),$$

and this is equivalentto $f(x, y) = x$ for every (x, y) $[0, 1]^2$. ∎

Let us see how Remark 5.55 and Theorem 5.56 apply to the Hammin g distance for fuzzy sets andtheIF-divergence ofHongand Kim.

Example 5.57 *Let usconsider the Hamming distance forfuzzy sets:*

$$l_{\mathrm{FS}}(A, B) = \sum_{i=1}^{n} |A(\omega_i) - B(\omega_i)|, \text{ for every } A,B \quad FS(\Omega).$$

Applying Theorem 5.50, we can build an IF-divergence from l_{FS}*:*

$$D_{\mathrm{IFS}}(A, B) = f\left(\sum_{i=1}^{n} |\mu_A(\omega_i) - \mu_B(\omega_i)|, \sum_{i=1}^{n} |\nu_A(\omega_i) - \nu_B(\omega_i)|\right),$$

andusingProposition5.43, wecanderive from D_{IFS} *another divergence* D *for fuzzy sets:*

$$D\,(A, B) = f\left(\sum_{i=1}^{n} |A(\omega_i) - B(\omega_i)|, |A(\omega_i) - B(\omega_i)|\right).$$

Then, $D(A, B) = l_{\mathrm{FS}}(A, B)$ *if and only if* $f(x, x) = x$ *. In particular* D *and* l_{FS} *are the same divergence iff* $f(x, y) = \frac{x+y}{2}$.

Consider now the IF-divergence D_C *defined byHong and Kimin Section5.1.3:*

$$D_C(A, B) = \frac{1}{2}\sum_{i=1}^{n} |\mu_A(\omega_i) - \mu_B(\omega_i)| + |\nu_A(\omega_i) - \nu_B(\omega_i)|.$$

Using Proposition 5.43 we can build a divergence forfuzzy sets:

$$D(A, B) = D_{\mathrm{IFS}}(A, B) = \sum_{i=1}^{n} |A(\omega_i) - B(\omega_i)| = l_{\mathrm{FS}}(A, B).$$

If we now apply Theorem 5.50, we can build other IF-divergence given by:

$$\begin{aligned} D_{\mathrm{IFS}}(A, B) &= f\,(D\,(A_1, B_1), D(A_2, B_2)) \\ &= f\left(\sum_{i=1}^{n} |\mu_A(\omega_i) - \mu_B(\omega_i)|, \sum_{i=1}^{n} |\nu_A(\omega_i) - \nu_B(\omega_i)|\right). \end{aligned}$$

Thus, we conclude that $D_{\mathrm{IFS}}(A, B) = D_C(A, B)$ *if and only if* $f(x, y) = \frac{x+y}{2}$.

Corollary 5.58 *Let D be a divergence for fuzzy sets. Then, the diagram:*

$$\begin{tikzcd} D \arrow[r, bend left, "5.57"] & D_{IFS} \arrow[l, bend left, "5.50"] \end{tikzcd}$$

commutes if and only if $f(x, y) = x$ *and*

$$D_{IFS}(A, B) = D(A_1, B_1), \text{ for every } A, B \quad IFSs(\Omega).$$

Proof On the one hand, from Theorem 5.56 we know that $f(x, y) = x$. Moreover, from Remark 5.55 the following equation must hold:

$$\begin{aligned} D_{IFS}(A, B) &= f(D_{IFS}(A_1, B_1), D_{IFS}(A_2, B_2)) = D_{IFS}(A_1, B_1) \\ &= f(D(A_1, B_1), D(A_2, B_2)) = D(A_1, B_1). \end{aligned}$$

Thus, for every A, B $IFSs(\Omega)$ it must hold that:

$$D_{IFS}(A, B) = D(A_1, B_1). \blacksquare$$

5.2 Connecting IVF-sets and imprecise probabilities

This section is devoted to investigate the relationship between IF-sets and Imprecise Probabilities. In fuzzy set theory, it is well known ([217]) that there exists a connection between fuzzy sets and possibility measures. In fact, given a normalized fuzzy set μ_A, it defines a possibility distribution with associated possibility measure Π defined by:

$$\Pi(B) = \sup_{x \; B} \mu_A(x).$$

Conversely given a possibility measure Π with associated possibility distribution π, it defines a fuzzy set with membership function π.

In this section, we shall assume first of all that the IVF-sets are defined on a probability space. Thus, any IVF-set defines a random set, and then the probabilistic information of the IVF-set can be summarized by means of the set of distributions of the measurable selections. In this framework, we investigate in which situations the probabilistic information can be equivalently represented by the set of probabilities that dominate the lower probability induced by the random interval, and the conditions under which the upper probability induced by the random interval is a possibility measure.

Afterwards, we shall investigate other possible relationships between IVF-sets and imprecise probabilities. For instance, we shall see that the definition of probability for IVF-set given by Grzegorzewski and Mrowka ([86]) becomes a particular case in our theory. We also investigate how a one-to-one relation could be defined between IVF-sets, p-boxes and clouds.

5.2.1 Probabilistic information of IVF-sets

In this section we shall assume that IVF-sets are defined on a probability space. Then, they define random sets. We investigate how the probabilistic information of a IVF-set can be summarized by means of Imprecise Probabilities.

Since formally IVF-sets and IF-sets are equivalent, as we saw in Section 2.3, we shall denote IVF-sets by:

$$\{[\mu_A(\omega), 1-\nu_A(\omega)] : \omega \in \Omega\},$$

where μ_A and ν_A refer the membership and non-membership degree of the associated IF-set.

IVFS as random intervals

As we mentioned in Section 2.3, an IVF-set can be regarded as a model for the imprecise knowledge about the membership function of a fuzzy set A, in the sense that for every ω in the possibility space Ω, its membership degree belongs to the interval $[\mu_A(\omega), 1-\nu_A(\omega)]$. Hence, we can equivalently represent the IVF-set I_A by means of a multi-valued mapping $\Gamma_A : \Omega \to \mathcal{P}([0,1])$, where

$$\Gamma_A(\omega) := [\mu_A(\omega), 1-\nu_A(\omega)]. \tag{5.5}$$

If the intuitionistic fuzzy set is defined on a probability space $(\Omega, \mathcal{A}, P)$, then the probabilistic information encoded by the multi-valued mapping Γ_A can be summarized by means of its lower and upper probabilities $\underline{P}_{\Gamma_A}, \overline{P}_{\Gamma_A}$. Recall that, from Equation (2.22), for any subset B in the Borel σ-field $\beta_{[0,1]}$, its lower and upper probabilities are given by

$$\underline{P}_{\Gamma_A}(B) := P(\{\omega : \Gamma_A(\omega) \subseteq B\})$$

and

$$\overline{P}_{\Gamma_A}(B) := P(\{\omega : \Gamma_A(\omega) \cap B \neq \emptyset\}).$$

We need to make two clarifications here: the first one is that the images of the multi-valued mapping Γ_A are non-empty, as a consequence of the restriction $\mu_A \leq 1-\nu_A$ in the definition of IVF-sets; the second is that, in order to be able to define the lower and upper probabilities $\underline{P}_{\Gamma_A}, \overline{P}_{\Gamma_A}$, the multi-valued mapping Γ_A needs to be *strongly measurable* ([88]), which in this case ([129]) means that the mappings

$$\mu_A, \nu_A : \Omega \to [0,1]$$

must be $\mathcal{A}-\beta_{[0,1]}-$measurable.

If we assume that the 'true' membership function imprecisely specified by means of the IVF-set is $A-\beta_{[0,1]}-$measurable, then it must belong to the set of measurable selections of Γ_A (see Equation (2.21)):

$$S(\Gamma_A) := \{\varphi : \Omega \to [0,1] \text{ measurable: } \varphi(\omega) \in [\mu_A(\omega), 1-\nu_A(\omega)]\ \forall \omega \in \Omega\},$$

and as a consequence the probability measure it induces will belong to the set

$$P(\Gamma_A) := \{P_\varphi : \varphi \in S(\Gamma_A)\}.$$

Any probability measure in $P(\Gamma_A)$ is bounded by the upper probability P_{Γ_A}, and as a consequence the set $P(\Gamma_A)$ is included in the set $M(P_{\Gamma_A})$ of probability measures that are dominated by P_{Γ_A}. As we have seen in Section 2.2.4, both sets are not equivalent in general; however, Proposition 2.45 shows several situations in which they coincide. Taking this result into account, we can establish the following conditions for the equality between the credal sets generated by an IVF-set.

Corollary 5.59 *Consider the initial space* $([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$ *and* $\Gamma_A : [0,1] \to \mathcal{P}([0,1])$ *defined as in Equation* (5.5). *Then, the equality* $M(P_{\Gamma_A}) = P(\Gamma_A)$ *holds under any of the following conditions:*

(a) The membership function μ_A *is increasing and the non-membership function* ν_A *is decreasing.*

(b) $\mu_A(\omega) = 0$ *for any* $\omega \in \Omega$.

(c) For any $\omega, \omega' \in \Omega$, *either* $\Gamma_A(\omega) \leq \Gamma_A(\omega')$ *or* $\Gamma_A(\omega) \geq \Gamma_A(\omega')$, *where* $[a_1, b_1] \leq [a_2, b_2]$ *if* $a_1 \leq a_2$ *and* $b_1 \leq b_2$.

The previous conditions can be interpreted as follows:

(a) The greater the value of ω, the more evidence supports that ω belongs to A.

(b) There is no evidence supporting that the elements belong to set A.

(c) The intervals associated with the elements are ordered. In particular, this holds when the hesitation is constant.

Proof

(a) Condition (3a) of Proposition 2.45 assures that $M(P_{\Gamma_A})$ and $P(\Gamma_A)$ coincide whenever the bounds of the random interval are increasing. In the particular case of IVF-set, this means that both μ_A and $1-\nu_A$ are increasing, or equivalently, that μ_A is increasing and ν_A is decreasing.

(b) Condition (3b) of Proposition 2.45 assures that $M(P_{\Gamma_A})$ and $P(\Gamma_A)$ coincide if the lower bound of the interval equals 0. In the case of IVF-sets, this means that $\mu_A = 0$.

(c) Condition (3c) of Proposition 2.45 assures that $M(P_{\Gamma_A})$ and $P(\Gamma_A)$ coincide if the bounds of the interval are strictly comonotone. In the case of IVF-sets, the bounds of the interval, μ_A and $1-\nu_A$, are comonotone if and only if $\Gamma_A(\omega) \geq \Gamma_A(\omega)$ or $\Gamma_A(\omega) \leq \Gamma_A(\omega)$ for any ω, ω: assume that μ_A and $1-\nu_A$ are comonotone, then $\mu_A(\omega) \geq \mu_A(\omega)$ if and only if $1-\nu_A(\omega) \geq 1-\nu_A(\omega)$ for every ω, ω. Thus:

- If $\mu_A(\omega) > \mu_A(\omega)$, then $1-\nu_A(\omega) > 1-\nu_A(\omega)$, so

$$\Gamma_A(\omega) = [\mu_A(\omega), 1-\nu_A(\omega)] > [\mu_A(\omega), 1-\nu_A(\omega)] = \Gamma_A(\omega).$$

- If $\mu_A(\omega) < \mu_A(\omega)$, then $1-\nu_A(\omega) < 1-\nu_A(\omega)$, so

$$\Gamma_A(\omega) = [\mu_A(\omega), 1-\nu_A(\omega)] < [\mu_A(\omega), 1-\nu_A(\omega)] = \Gamma_A(\omega).$$

On the other hand, assume that either $\Gamma_A(\omega) \geq \Gamma_A(\omega)$ or $\Gamma_A(\omega) \leq \Gamma_A(\omega)$ for any ω, ω. Then:

$$\begin{array}{ll} \Gamma_A(\omega) \geq \Gamma_A(\omega) & \mu_A(\omega) \geq \mu_A(\omega) \text{ and } 1-\nu_A(\omega) \geq 1-\nu_A(\omega) \\ \Gamma_A(\omega) \leq \Gamma_A(\omega) & \mu_A(\omega) \leq \mu_A(\omega) \text{ and } 1-\nu_A(\omega) \leq 1-\nu_A(\omega) \end{array}$$

and from this we deduce that μ_A and $1-\nu_A$ are comonotone. ∎

On the other hand, [129, Example 3.3] shows that the equality $P(\Gamma) = M(P_\Gamma)$ does not necessarily hold for all the random closed intervals, even when the initial probability space is non-atomic: it suffices to consider $(\Omega, \mathcal{A}, P) = ([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$ and $\Gamma : [0,1] \to \mathcal{P}(\mathbb{R})$ given by

$$\Gamma(\omega) = [-\omega, \omega] \quad \omega \in [0,1].$$

It is easy to adapt the example to our context and deduce that there are intuitionistic fuzzy sets where the information about the membership function is not completely determined by the upper probability P_{Γ_A}: it would suffice to take $\Gamma_A : [0,1] \to \mathcal{P}([0,1])$ given by

$$\Gamma_A(\omega) = \left[0.5 - \frac{\omega}{2}, 0.5 + \frac{\omega}{2}\right] \quad \omega \in [0,1], \tag{5.6}$$

that is, to consider the IVF-set such that the membership and non-membership functions of its associated IF-set coincide and take the value $\frac{1-\omega}{2}$.

We have seen in Proposition 2.47 that the upper probability associated with a random set is a possibility measure if and only if the images of Γ are nested except for a null subset. In the particular case of the random closed intervals associated with an IVF-set, we deduce the following:

Corollary 5.60 *Let* $\Gamma_A : \Omega \to \mathcal{P}([0,1])$ *be the random set defined in the probability space* $(\Omega, \mathcal{A}, P)$ *by Equation* (5.5). *Then,* P_Γ *is possibility measure ifand onlyif there exists some* $N \quad \Omega$ *null such that* μ_A *and* ν_A *are comonotoneon* $\Omega \setminus N$.

Pro of Assume that Γ_A is a possibility measure. Then, by Prop osition 2.47, the re is anull set N such that $\Gamma_A(\omega_1) \quad \Gamma_A(\omega_2)$ or $\Gamma_A(\omega_2) \quad \Gamma_A(\omega_1)$ for any $\omega_1, \omega_2 \quad \Omega \setminus N$. Consider $\omega_1, \omega_2 \quad \Omega \setminus N$, it holds that:

$$
\begin{aligned}
\Gamma_A(\omega_1) \quad \Gamma_A(\omega_2) \quad & [\mu_A(\omega_1), 1-\nu_A(\omega_1)] \quad [\mu_A(\omega_2), 1-\nu_A(\omega_2)] \\
& \mu_A(\omega_1) \geq \mu_A(\omega_2) \text{ and } 1-\nu_A(\omega_1) \leq 1-\nu_A(\omega_2) \\
& \mu_A(\omega_1) \geq \mu_A(\omega_2) \text{ and } \nu_A(\omega_1) \geq \nu_A(\omega_2) \\
\Gamma_A(\omega_2) \quad \Gamma_A(\omega_1) \quad & [\mu_A(\omega_2), 1-\nu_A(\omega_2)] \quad [\mu_A(\omega_1), 1-\nu_A(\omega_1)] \\
& \mu_A(\omega_2) \geq \mu_A(\omega_1) \text{ and } 1-\nu_A(\omega_2) \leq 1-\nu_A(\omega_1) \\
& \mu_A(\omega_2) \geq \mu_A(\omega_1) \text{ and } \nu_A(\omega_2) \geq \nu_A(\omega_1).
\end{aligned}
$$

Then, μ_A and ν_A arecomonotone on $\Omega \setminus N$.

Conversely, as sume that μ_A and ν_A arecomonotone on $\Omega \setminus N$.

$$
\begin{aligned}
&\text{If } \mu_A(\omega_1) \leq \mu_A(\omega_2) \quad \nu_A(\omega_1) \leq \nu_A(\omega_2) \\
&\qquad [\mu_A(\omega_1), 1-\nu_A(\omega_1)] \quad [\mu_A(\omega_2), 1-\nu_A(\omega_2)] \quad \Gamma_A(\omega_2) \quad \Gamma_A(\omega_1). \\
&\text{If } \mu_A(\omega_2) \leq \mu_A(\omega_1) \quad \nu_A(\omega_2) \leq \nu_A(\omega_1) \\
&\qquad [\mu_A(\omega_2), 1-\nu_A(\omega_2)] \quad [\mu_A(\omega_1), 1-\nu_A(\omega_1)] \quad \Gamma_A(\omega_1) \quad \Gamma_A(\omega_2).
\end{aligned}
$$

P-b ox induced by a IVF-set

The lower and upp er probabilities $\underline{P}_{\Gamma_A}, \overline{P}_{\Gamma_A}$ summarize the probabilistic information ab out the probability distribution of the memb ership function of th e IVF-set A. If in particular we want to summarise the information ab out the distribution function of this variable, we must use the lower and upp er distribution functions:

$$\underline{F}_A, \overline{F}_A : \Omega \to [0, 1],$$

where

$$\underline{F}_A(x) := \underline{P}_{\Gamma_A}([0, x]) = P(\{\omega : 1-\nu_A(\omega) \leq x\}) = P_{\nu_A}([1-x, 1]) \tag{5.7}$$

and

$$\overline{F}_A(x) := \overline{P}_{\Gamma_A}([0, x]) = P(\{\omega : \mu_A(\omega) \leq x\}) = P_{\mu_A}([0, x]). \tag{5.8}$$

When Ω is an ordered space(for instance if $\Omega = [0, 1]$), the lower and upp er distribution functions $\underline{F}_A, \overline{F}_A$ can be used to determinea p-b ox. In that case, we shall refer to $(\underline{F}_A, \overline{F}_A)$ asthe *p-box on* Ω *associated with the intuitionistic fuzzy set* A.

The lower and upper distribution functions also determine a set of probability measures:

$$M(\underline{F}_A, \overline{F}_A) := \{Q : \beta_{[0,1]} \to [0,1] : \underline{F}_A(x) \le F_Q(x) \le \overline{F}_A(x) \quad \forall x \in [0,1]\},$$

where F_Q is the distribution function associated with the probability measure Q. It is immediate to see that the set $M(\underline{F}_A, \overline{F}_A)$ includes $M(P_{\Gamma_A})$. However, the two sets do not coincide in general, and as a consequence the use of the lower and upper distribution functions may produce a loss of information, as we can see in the following example.

Example 5.61 *Consider the random set of Equation* (5.6), *defined on* $([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$ *by* $\Gamma_A(\omega) = [0.5 - \frac{\omega}{2}, 0.5 + \frac{\omega}{2}]$. *Using Equation* (2.23), *we already know that the credal set* $M(P_{\Gamma_A})$ *is given by:*

$$M(P_{\Gamma_A}) = \{P \text{ probability} \mid P_{\Gamma_A}(B) \le P(B) \le P^{\Gamma_A}(B) \text{ for any } B\}.$$

Let us now compute the form of the set $M(\underline{F}_A, \overline{F}_A)$:

$$\begin{aligned}
\underline{F}_A(x) &= P_{\Gamma_A}([0,x]) = P(\{\omega \in [0,1] : \Gamma(\omega) \subseteq [0,x]\}) \\
&= P(\{\omega \in [0,1] : \Gamma(\omega) = [0.5 - \tfrac{\omega}{2}, 0.5 + \tfrac{\omega}{2}] \subseteq [0,x]\}) \\
&= P(\{\omega \in [0,1] : \omega \in [-1, 2x-1]\}) \\
&= P(\{\omega \in [0,1] : \omega \in [0, 2x-1]\}) \\
&= \begin{cases} 0 & \text{if } x \le \frac{1}{2}. \\ 2x - 1 & \text{otherwise.} \end{cases} \\
\overline{F}_A(x) &= P^{\Gamma_A}([0,x]) = P(\{\omega \in [0,1] : \Gamma(\omega) \cap [0,x] \neq \emptyset\}) \\
&= P(\{\omega \in [0,1] : [0.5 - \tfrac{\omega}{2}, 0.5 + \tfrac{\omega}{2}] \cap [0,x] \neq \emptyset\}) \\
&= \begin{cases} 2x & \text{if } x \le \frac{1}{2}. \\ 1 & \text{otherwise.} \end{cases}
\end{aligned}$$

Thus, the set $M(\underline{F}_A, \overline{F}_A)$ *is formed by the probabilities whose associated cumulative distribution function is bounded by* $\underline{F}_A$ *and* $\overline{F}_A$.

Consider now the probability distribution associated with the cumulative distribution function F *defined by:*

$$F(x) = \begin{cases} \overline{F}_A(x) & \text{if } x \le \frac{1}{4}. \\ \frac{1}{2} & \text{if } x \in \left(\frac{1}{4}, \frac{3}{4}\right]. \\ \underline{F}_A(x) & \text{if } x > \frac{3}{4}. \end{cases}$$

Its associated probability, P_F, *belongs to* $M(\underline{F}_A, \overline{F}_A)$. *Now, let us check that* P_F *does not belong to* $M(P_{\Gamma_A})$. *For this aim, note that:*

$$\begin{aligned}
P_{\Gamma_A}\left(\left(\tfrac{1}{4}, \tfrac{3}{4}\right)\right) &= P\left(\left\{\omega \in [0,1] : \Gamma(\omega) \subseteq \left(\tfrac{1}{4}, \tfrac{3}{4}\right)\right\}\right) \\
&= P\left(\left\{\omega \in [0,1] : [0.5 - \tfrac{\omega}{2}, 0.5 + \tfrac{\omega}{2}] \subseteq \left(\tfrac{1}{4}, \tfrac{3}{4}\right)\right\}\right) \\
&= P\left(\left\{\omega \in [0,1] : \omega \in \left[0, \tfrac{1}{2}\right)\right\}\right) = \tfrac{1}{2}.
\end{aligned}$$

This means that every probability P in $M(P_{\Gamma_A})$ must hold that $P\ \frac{1}{4}, \frac{3}{4}\ \geq \frac{1}{2}$. However, $P_F\ \frac{1}{4}, \frac{3}{4} = 0$, and consequently $P_F \ / \ M(P_{\Gamma_A})$.

We conclude that $M(\underline{F}_A, \overline{F}_A)\quad M(P_{\Gamma_A})$.

Nevertheless, there are non-trivial situations in which both sets coincide.

Example 5.62 *Consider the initial space $([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$ and the random set Γ_A defined from the IF-set I_A by:*

$$\Gamma_A(\omega) = \begin{cases} \{\omega\} & \text{if } \omega \quad 0, \frac{1}{4} \quad \frac{3}{4}, 1. \\ \frac{1}{4}, \frac{3}{4} & \text{otherwise.} \end{cases}$$

Thus, the membership and non-membership functions are given by:

$$\mu_A(\omega) = \begin{cases} \omega & \text{if } \omega \quad 0, \frac{1}{4} \quad \frac{3}{4}, 1 \\ \frac{1}{4} & \text{otherwise,} \end{cases}$$

and

$$\nu_A(\omega) = \begin{cases} 1-\omega & \text{if } \omega \quad 0, \frac{1}{4} \quad \frac{3}{4}, 1. \\ \frac{1}{4} & \text{otherwise.} \end{cases}$$

Then, the lower and upper cdfs $\underline{F}_A$ and $\overline{F}_A$ are given by:

$$\underline{F}_A(x) = \begin{cases} x & \text{if } x \quad 0, \frac{1}{4}, \\ \frac{1}{4} & \text{if } x \quad \frac{1}{4}, \frac{3}{4}, \\ x & \text{if } x \quad \frac{3}{4}, 1, \end{cases} \quad \text{and } \overline{F}_A(x) = \begin{cases} x & \text{if } x \quad 0, \frac{1}{4}. \\ \frac{3}{4} & \text{if } x \quad \frac{1}{4}, \frac{3}{4}. \\ x & \text{if } x \quad \frac{3}{4}, 1. \end{cases}$$

We know that $M(\underline{F}_A, \overline{F}_A)\quad M(P_{\Gamma_A})$. Let us now see that for every probability P such that $\underline{F}_A \leq F_P \leq \overline{F}_A$, $P\quad M(P_{\Gamma_A})$. Let P be one such probability, and let F_P denote its associated cumulative distribution function. Consider now the measurable map $U(\omega) := F_P^{-1]}(\omega)$, where $F_P^{-1]}$ denotes the pseudo-inverse of the cumulative distribution function F_P. It trivially holds that $U\quad S(\Gamma_A)$, and consequently $P_U\quad P(\Gamma_A)\quad M(P_{\Gamma_A})$. On the other hand, since $F_P^{-1]}(\omega) \leq x$ if and only if $\omega\quad [0, F_P(x)]$, F_U and F_P coincide:

$$\begin{aligned} F_U(x) &= P(\{\omega \quad [0,1] \mid U(\omega) \leq x\}) = P(\{\omega \quad [0,1] \mid F_P^{-1]}(\omega) \leq x\}) \\ &= P(\{\omega \quad [0,1] \mid \omega \leq F_P(x)\}) = P([0, F_P(x)]) = F_P(x). \end{aligned}$$

Thus, $P = P_U$, and consequently $P\quad P(\Gamma_A)\quad M(P_{\Gamma_A})$.

The following result gives a sufficient condition for the equality between $M(\underline{F}_A, \overline{F}_A)$ and $M(P_{\Gamma_A})$:

Proposition 5.63 *If the initial space is* $([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$ *and the random interval is an IVF-set as in Equation* (5.5)*, where* $\mu_A(x) = 0$ *for every* x*, then* $M(\underline{F}_A, \overline{F}_A) = M(P_{\Gamma_A})$.

Proof Assume there is a probability $P \in M(\underline{F}_A, \overline{F}_A)$ such that for some measurable B it satisfies $P(B) \notin [P_{\Gamma_A}(B), P_{\Gamma_A}(B)]$. We consider two cases: $0 \in B$ and $0 \notin B$.

$0 \notin B$**:** When $0 \notin B$, it holds that $P_{\Gamma_A}(B) = 0$:

$$P_{\Gamma_A}(B) = P(\{\omega \mid \Gamma(\omega) \subseteq B\}) = P(\{\omega \mid [0, \mu_A(\omega)] \subseteq B\}) = 0,$$

since $0 \in \Gamma_A(\omega) \setminus B$ for any ω. Then, it holds that $P(B) > P_{\Gamma_A}(B)$. In addition, $P_{\Gamma_A}(B) = 1 - P_{\Gamma_A}(B^c)$, and consequently $P_{\Gamma_A}(B^c)$ must be strictly positive (otherwise $P(B) > P_{\Gamma_A}(B) = 1$ and a contradiction arises). Thus, there exists an interval $[0, x] \subseteq B^c$. Let $\varepsilon = \sup\{x : [0, x] \subseteq B^c\}$, and consider two cases:

- Assume that $\varepsilon = \max\{x : [0, x] \subseteq B^c\}$. Then, since $(\varepsilon, 1] \supseteq B$, it holds that:

  $$P(B) \leq P((\varepsilon, 1]) = 1 - F_P(\varepsilon),$$

  and consequently:

  $$1 - F_P(\varepsilon) \geq P(B) > P_{\Gamma_A}(B) = 1 - P_{\Gamma_A}(B^c).$$

  But:

  $$P_{\Gamma_A}(B^c) = P(\{\omega \mid \Gamma_A(\omega) \subseteq B^c\}) = P(\{\omega \mid \Gamma_A(\omega) \subseteq [0, \varepsilon]\}) = \underline{F}_A(\varepsilon).$$

  Thus:

  $$1 - F_P(\varepsilon) > 1 - P_{\Gamma_A}(B^c) = 1 - \underline{F}_A(\varepsilon) \Rightarrow \underline{F}_A(\varepsilon) > F(\varepsilon),$$

  and a contradiction arises since $P \notin M(\underline{F}_A, \overline{F}_A)$.
- Assume that $\varepsilon = \max\{x : [0, x] \subseteq B^x\}$. Then:

  $$P_{\Gamma_A}(B^c) = P(\{\omega \mid \Gamma_A(\omega) \subseteq B^c\}) = P(\{\omega \mid \Gamma_A(\omega) \subseteq [0, \varepsilon)\}) = P_{\Gamma_A}([0, \varepsilon)).$$

  Moreover:

  $$[0, \varepsilon) \subseteq B^c \Rightarrow [\varepsilon, 1] \supseteq B \Rightarrow P([\varepsilon, 1]) \geq P(B).$$

  Thus, it holds that

  $$P([\varepsilon, 1]) \geq P(B) > 1 - P_{\Gamma_A}([0, \varepsilon)) = P_{\Gamma_A}([\varepsilon, 1]).$$

  However, note that $F_P(t) \geq \underline{F}_A(t) = F_{1-\nu_A}(t)$ for any t, and:

  $$P([\varepsilon, 1]) = 1 - F_P(t^-) \leq 1 - \underline{F}_A(t^-) = P_{\Gamma_A}([\varepsilon, 1]),$$

  a contradiction.

$0 \notin B$**:** Note that, since $0 \notin B$, $P_{\Gamma_A}(B) = 1$:

$$P_{\Gamma_A}(B) = P(\{\omega \mid \Gamma(\omega) \cap B = \emptyset\}) \geq P(\{\omega \mid \Gamma(\omega) \cap \{0\} = \emptyset\}) = 1.$$

Then $P(B) < P_{\Gamma_A}(B)$. Since $P_{\Gamma_A}(B) > 0$, there exists $[0, x] \subseteq B$. Define $\varepsilon = \sup\{x : [0, x] \subseteq B\}$ and consider two cases:

- Assume that $\varepsilon = \max\{x : [0, x] \subseteq B\}$. Then, $P(B) \geq P([0, \varepsilon]) = F_P(\varepsilon)$. However:

$$P_{\Gamma_A}(B) = P(\{\omega \mid \Gamma_A(\omega) \subseteq B\}) = P(\{\omega \mid \Gamma_A(\omega) \subseteq [0, \varepsilon]\})$$
$$= \overline{F}_A(\varepsilon) \leq F_P(\varepsilon) \leq P(B),$$

  a contradic tion, b ecause we had assumed that $P_{\Gamma_A}(B) > P(B)$.

- Assume that $\varepsilon = \max\{x : [0, x] \subseteq B\}$. Then $P(B) \geq P([0, \varepsilon))$. Moreover,

$$P_{\Gamma_A}(B) = P(\{\omega \mid \Gamma A(\omega) \subseteq B\}) = P(\{\omega \mid \Gamma A(\omega) \subseteq [0, \varepsilon)\})$$
$$= \overline{F}_A(\varepsilon^-) = F_{1-\nu_A}(\varepsilon^-) \leq F_P(\varepsilon^-) = P([0, \varepsilon)) \leq P(B).$$

  This contradicts the assumption of $P_{\Gamma_A}(B) > P(B)$. ∎

Another suffic ient condition for the equality b etween $M(P_{\Gamma_A})$ and $M(\underline{F}_A, \overline{F}_A)$ is the strict comonotonicity between μ_A and $1 - \nu_A$, that, as we have seen in Corollary 5.59, is equivalent to the exis tence of a total order b etween the intervals $[\mu_A(\omega), 1 - \nu_A(\omega)]$.

Prop osition 5.64 *If the initial space is* $([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$ *and the random interval is given by an IF-set as in Equation (5.5), where* $\Gamma_A(\omega) \leq \Gamma_A(\omega)$ *or* $\Gamma_A(\omega) \geq \Gamma_A(\omega)$ *for any* $\omega, \omega \in \Omega$*, then* $M(\underline{F}_A, \overline{F}_A) = M(P_{\Gamma_A})$.

Pro of In [129, Theorem 4.5] it is proven that when the random interval is defined on $([0, 1], \beta_{[0,1]}, \lambda_{[0,1]})$ and its bounds are strictly comonotone, then it is p ossibl e to define the random interval $\Gamma : [0, 1] \to \mathcal{P}([0, 1])$ by:

$$\Gamma(\omega) := [U(\omega), V(\omega)],$$

where U and V denote th e quantile functions of the lower and upp er bounds of Γ_A, resp ectively, that are defined by:

$$U(\omega) = \inf\{x \in \mathbb{R} : \omega \leq \underline{F}(x)\} \text{ and } V(\omega) = \inf\{x \in \mathbb{R} : \omega \leq \overline{F}(x)\}.$$

This random interval satisfies $P_\Gamma = P_{\Gamma_A}$, and consequently $M(P_\Gamma) = M(P_{\Gamma_A})$ and $M(\underline{F}, \overline{F}) = M(\underline{F}_A, \overline{F}_A)$. Then, in order to prove the equality $M(P_{\Gamma_A}) = M(\underline{F}_A, \overline{F}_A)$ it is sufficient to establish the equality between $M(P_\Gamma) = M(\underline{F}, \overline{F})$.

Consider now a probability $P \in M(\underline{F}, \overline{F})$, and define W as the quantile function of F_P. Since $\underline{F} \leq F_P \leq \overline{F}$, $W(\omega)$ is bounded by $U(\omega)$ and $V(\omega)$ for any $\omega \in [0, 1]$. Then, W

is a measurable selection of Γ, and its induced probability P_W belongs to $P(\Gamma)$. Moreover, since $P(\Gamma) = M(P_\Gamma)$, P_W also belongs to $M(P_\Gamma)$.

Thus, $M(P_\Gamma) = M(\underline{F}, \overline{F})$, and therefore $M(P_{\Gamma_A}) = M(\underline{F}_A, \overline{F}_A)$. ∎

One particular situation where the previous result holds is when μ_A is strictly increasing, ν_A is strictly decreasing and $\mu_A(\omega) = \mu_A(\omega')$ if and only if $\nu_A(\omega) = \nu_A(\omega')$.

Finally, we are going to see that the equality between both credal sets also holds when the bounds of the interval are increasing.

Proposition 5.65 *If the initial space is* $([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$ *and the random interval is given by an IF-set as in Equation* (5.5), *where* μ_A *is increasing and* ν_A *is decreasing, then* $M(\underline{F}_A, \overline{F}_A) = M(P_{\Gamma_A})$.

Proof Let P be a probability in $M(\underline{F}_A, \overline{F}_A)$, and we are going to see that there is a measurable selection V such that $P_V = P$, and therefore $M(\underline{F}_A, \overline{F}_A) \subseteq P(\Gamma_A) \subseteq M(P_{\Gamma_A})$. Since μ_A is increasing, there is a countable number of elements $\omega \in (0,1)$ such that $\mu_A(\omega) > \sup_{\omega' < \omega} \mu_A(\omega')$. Denote this set by N, and consider the function $V : [0,1] \to \mathbb{R}$ defined by:

$$V(\omega) = \begin{cases} \inf\{y : \omega \leq P((-\infty, y])\} & \text{if } \omega \in (0,1) \setminus N. \\ \mu_A(\omega) & \text{otherwise.} \end{cases}$$

Following the same steps than in [129, Proposition 4.1], this function V can be proved to be a measurable selection of Γ_A such that $P_V = P$. Then, we conclude that $M(\underline{F}_A, \overline{F}_A) \subseteq P(\Gamma_A) \subseteq M(P_{\Gamma_A})$, and then we conclude that both credal sets coincide. ∎

These results allow us state a number of sufficient conditions for the equality between the three sets of probabilities $P(\Gamma_A)$, $M(P_{\Gamma_A})$ and $M(\underline{F}_A, \overline{F}_A)$.

Corollary 5.66 *Consider the initial space is* $([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$ *and the random interval* Γ_A *given by an IVF-set as in Equation* (5.5). *Then, the equalities* $P(\Gamma_A) = M(P_{\Gamma_A}) = M(\underline{F}_A, \overline{F}_A)$ *hold if one of the following conditions is satisfied:*

- μ_A *is increasing and* ν_A *is decreasing.*
- $\mu_A(\omega) = 0$ *for any* $\omega \in [0,1]$.
- μ_A *and* $1 - \nu_A$ *are strictly comonotone, or equivalently, if* $\Gamma_A(\omega) \leq \Gamma_A(\omega')$ *or* $\Gamma_A(\omega') \leq \Gamma_A(\omega)$ *for any* $\omega, \omega' \in [0,1]$.

We have seen sufficient conditions under which the p-box defined from the random interval Γ_A contains the same information than the set of measurable selections. Conversely, there are situations in which, given a p-box, it is possible to define a random interval Γ_A whose associated p-box coincides with the previous one, and that the probabilistic information given by the p-box is the same that the information given by the set of measurable selections.

Proposition 5.67 *Consider a p-box $(\underline{F}, \overline{F})$ defined on $[0,1]$ such that both $\underline{F}$ and $\overline{F}$ are right-continuous. Then it is possible to define a random interval $\Gamma : [0,1] \to \mathcal{P}([0,1])$ whose associated p-box is $(\underline{F}, \overline{F})$. In addition, if either $\underline{F}$ and $\overline{F}$ are strictly comonotone or $\overline{F}(x) = 1$, then the random interval Γ satisfies $P(\Gamma) = M(\underline{F}, \overline{F})$.*

Proof Proposition 2.45 assures that $P(\Gamma) = M(P_{\Gamma_A})$. Given the p-box $(\underline{F}, \overline{F})$, define the random interval $\Gamma_A(\omega) = [U(\omega), V(\omega)]$, where U and V are the quantile functions of $\underline{F}$ and $\overline{F}$, respectively. Then, the p-box associated with Γ_A is given by:

$$\underline{F}_A(t) = F_V(t) = P(\{\omega \in [0,1] \mid V(\omega) \leq t\}) = \underline{F}(t).$$
$$\overline{F}_A(t) = F_U(t) = P(\{\omega \in [0,1] \mid U(\omega) \leq t\}) = \overline{F}(t).$$

Since $\underline{F}$ and $\overline{F}$ are right-continuous, U and V are random variables because their cumulative distribution functions are right-continuous. Assume now that $\underline{F}$ and $\overline{F}$ are strictly comonotone. Then, U and V are also strictly comonotone, and following Proposition 5.64, the credal set $P(\Gamma_A)$ coincides with the credal set $M(\underline{F}, \overline{F})$.

Assume that $\overline{F}(x) = 1$. Then, $U = 0$ almost surely. Applying Proposition 5.63, $P(\Gamma_A) = M(\underline{F}, \overline{F})$. ■

In Corollary 5.60 we have seen that the upper probability induced by the random set Γ_A defined from an IF-set I_A is a possibility measure if and only if μ_A and ν_A are strictly comonotone on the complementary of a null set. In [199], the following result is proved:

Proposition 5.68 ([199, Corollary 17]) *Assume that $\Omega/\!\sim$ is order complete and let $(\underline{F}, \overline{F})$ be a p-box. Let $\underline{P}_{(\underline{F},\overline{F})}$ denote the lower probability associated with $(\underline{F}, \overline{F})$ by means of Equation (2.17). Then the natural extension of $\underline{P}_{(\underline{F},\overline{F})}$ is a possibility measure if and only if either*

(L1) $\underline{F}$ is 0–1 valued,

(L2) $\overline{F}(x) = \overline{F}(x^-)$ for all $x \in \Omega$ that have no immediate predecessor, and

(L3) $\{x \in \Omega \cup \{0^-\} : \underline{F}(x) = 1\}$ has a minimum, where 0^- is a minimum element on Ω,

or

(U1) $\overline{F}$ *is 0–1 valued,*

(U2) $\overline{F}(x) = \overline{F}(x^+)$ *for all* $x \in \Omega$ *that have no immediate successor, and*

(U3) $\{x \in \Omega \cup \{0^-\} : \overline{F}(x) = 0\}$ *has a maximum.*

In our context, when the initial space is $[0,1]$, no element in such interval has immediate predecessor or successor. Assume now that the p-box $(\underline{F}_A, \overline{F}_A)$ defined from the random interval Γ_A as in Equations (5.7) and (5.8) is a possibility measure. Note that since $\underline{F}_A$ and $\overline{F}_A$ are right-continuous, $(U2)$ becomes trivial. On the one hand, assume that $\underline{F}_A$ is 0–1 valued. Then, there exists t such that $\underline{F}(t) = 1$ for any $t \geq t$ and $\underline{F}(t) = 0$ for any $t < t$, and by $(L3)$ it is left-continuous. Equivalently:

$$\underline{F}(t) = P(\{\omega \in [0,1] : \Gamma_A(\omega) \subseteq [0,t]\}) = 1 \quad \text{for any } t \geq t.$$
$$\underline{F}(t) = P(\{\omega \in [0,1] : \Gamma_A(\omega) \subseteq [0,t]\}) = 0 \quad \text{for any } t < t.$$

Then, $1 - \nu_A(\omega) = t$ for every $\omega \in [0,1] \setminus N$ for some null set N on $\beta_{[0,1]}$. On the other hand, assume that $\overline{F}_A$ is 0–1 valued. Then, there exists t such that $\overline{F}(t) = 1$ for any $t > t$ and $\overline{F}(t) = 0$ for any $t \leq t$, and by $(U2)$ it is right-continuous. Equivalently:

$$\overline{F}(t) = P(\{\omega \in [0,1] : \Gamma_A(\omega) \cap [0,t] \neq \emptyset\}) = 1 \quad \text{for any } t \geq t.$$
$$\overline{F}(t) = P(\{\omega \in [0,1] : \Gamma_A(\omega) \cap [0,t] \neq \emptyset\}) = 0 \quad \text{for any } t < t.$$

Thus, $\mu_A(\omega) = t$ for every $\omega \in [0,1] \setminus N$ for some null set N on $\beta_{[0,1]}$. We deduce that:

Proposition 5.69 *Consider the initial space* $([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$ *and the random interval* Γ_A *defined from the IVF-set* I_A. *Consider the p-box* $(\underline{F}_A, \overline{F}_A)$ *defined in Equations* (5.7) *and* (5.8). *If* $(\underline{F}_A, \overline{F}_A)$ *defines a possibility measure, then there is a null set* N *on* $\beta_{[0,1]}$ *and* t *such that either* $1 - \nu_A(\omega) = t$ *for any* $\omega \in [0,1] \setminus N$ *or* $\mu_A(\omega) = t$ *for any* $\omega \in [0,1] \setminus N$. *In such a case,* $P(\Gamma_A) = M(P_{\Gamma_A}) = M(\underline{F}_A, \overline{F}_A)$.

A non-measurable approach

The previous developments assume that the intuitionistic fuzzy set is defined on a probability space and that the functions μ_A, ν_A are measurable with respect to the σ-field we have on this space and the Borel σ-field on $[0,1]$. Although this is a standard assumption when considering the probabilities associated with fuzzy events, it is arguably done for mathematical convenience only. In this section, we present an alternative approach where we get rid of the measurability assumptions by means of finitely additive probabilities. This allows us to make a clearer link with p-boxes, by means of Walley's notion of natural extension introduced in Definition 2.32.

Consider thus a intuitionistic fuzzy set A defined on a space Ω. If this set is determined by the functions μ_A, ν_A, we can represent it by means of the multi-valued mapping

$\Gamma_A : \Omega \to [0, 1]$ given by $\Gamma_A(\omega) = [\mu_A(\omega), 1 - \nu_A(\omega)]$. Note that we are not assuming anymore that this multi-valued mapping is strongly measurable, and now our information about the "true" membership function would be given by the set of functions

$$\{\varphi : \Omega \to [0, 1] : \mu_A(\omega) \leq \varphi(\omega) \leq 1 - \nu_A(\omega)\}.$$

Now, if we do not assume the measurability of μ_A, ν_A and consider then the field $\mathcal{P}(\Omega)$ of all events in the initial space, we may not be able to model our uncertainty by means of a σ-additive probability measure. However, we can do so by means of a finitely additive probability measure P or more generally by means of an imprecise probability model [205]. Moreover, the notions of lower and upper probabilities can be generalized to that case [132]. If for instance we consider a finitely additive probability P on $\mathcal{P}(\Omega)$, then by an analogous reasoning to that in Section 5.2.1 we obtain that

$$P_\varphi(C) \in [\underline{P}_{\Gamma_A}(C), \overline{P}_{\Gamma_A}(C)] \quad \forall C \subseteq [0, 1],$$

where $\overline{P}_{\Gamma_A}$ is the completely alternating upper probability given by

$$\overline{P}_{\Gamma_A}(C) = P(\{\omega : \Gamma_A(\omega) \cap C \neq \emptyset\})$$

and its conjugate $\underline{P}_{\Gamma_A}$ is the completely monotone lower probability given by

$$\underline{P}_{\Gamma_A}(C) = P(\{\omega : \emptyset \neq \Gamma_A(\omega) \subseteq C\})$$

for every $C \subseteq [0, 1]$. Then the information about P_φ is given by the set of finitely additive probabilities dominated by $\overline{P}_{\Gamma_A}$, and we do not need to make the distinction between $P(\Gamma_A)$ and $M(\overline{P}_{\Gamma_A})$ as in Section 5.2.1.

The associated p-box is given now by the set of finitely additive distribution functions (that is, monotone and normalized) that lie between $\underline{F}_A$ and $\overline{F}_A$, where again $\underline{F}_A, \overline{F}_A$ are given by Equations (5.7) and (5.8), respectively.

This set is equivalent to the set of associated finitely additive probability measures that can be determined by natural extension. This can be determined in the following way ([198]): if we denote by $\mathcal{H}$ the field of subsets of $[0, 1]$ generated by the sets $\{[0, x], (x, 1] : x \in [0, 1]\}$, then any set $B \in \mathcal{H}$ is of the form

$$B := [0, x_1] \cup (x_2, x_3] \cup \ldots \cup (x_{2n}, x_{2n+1}]$$

or

$$B := (x_1, x_2] \cup \ldots \cup (x_{2n}, x_{2n+1}]$$

for some $n \in \mathbb{N}, x_1 < x_2 < \cdots < x_n \in [0, 1]$. It holds that

$$\underline{E}_{\underline{F},\overline{F}}([0, x_1] \cup (x_2, x_3] \cup \ldots \cup (x_n, 1]) = \underline{F}_A(x_1) + \sum_{i=1}^{n} \max\{0, \underline{F}_A(x_{2i+1}) - \overline{F}_A(x_{2i})\}$$

and

$$E_{\underline{F},\overline{F}}((x_1,x_2]\cup\dots\cup(x_{2n},x_{2n+1}]) = \sum_{i=0}^{n} \max\{0, \underline{F}_A(x_{2i+1}) - \overline{F}_A(x_{2i})\} \tag{5.9}$$

and if we consider any $C\subseteq[0,1]$ then

$$E_{\underline{F},\overline{F}}(C) = \sup_{B\subseteq C, B\in\mathcal{H}} E_{\underline{F},\overline{F}}(B).$$

The upper probability $\overline{P}_{\Gamma_A}$ is determined by P_{Γ_A} using conjugacy.

It can be easily seen that P_{Γ_A} and the natural extension of the p-box $E_{\underline{F},\overline{F}}$ do not coincide in general, even in sets of the form $(x_1,x_2]$:

Example 5.70 *Consider the random interval of Example 5.61. We already know that* $P_{\Gamma_A}\left(\left(\frac14,\frac34\right]\right)=\frac12$. *Similarly, it can be proved that* $P_{\Gamma_A}\left(\left(\frac14,\frac34\right]\right)=\frac12$. *Now, let us use Equation* (5.9) *to compute* $E_{\underline{F},\overline{F}}\left(\left(\frac14,\frac34\right]\right)$:

$$E_{\underline{F},\overline{F}}\left(\left(\frac14,\frac34\right]\right)=\max\left\{0,\underline{F}_A\left(\frac34\right)-\overline{F}_A\left(\frac14\right)\right\}=\max\left\{0,\frac12-\frac12\right\}=0.$$

We conclude that, in general, P_{Γ_A} *and* $E_{\underline{F},\overline{F}}$ *do not coincide even in sets of the form* $(x_1,x_2]$.

Our next example shows that P_{Γ_A} and $E_{\underline{F},\overline{F}}$ do not coincide neither when the bounds of the random interval are increasing.

Example 5.71 *Consider the random interval defined by:*

$$\Gamma_A(\omega)=\begin{cases}[\omega,2\omega] & \text{if } \omega\in\left[0,\frac13\right].\\ \left[\frac13,\frac23\right] & \text{if } \omega\in\left(\frac13,\frac23\right).\\ [2\omega-1,\omega] & \text{otherwise.}\end{cases}$$

The bounds of its associated p-box are defined by:

$$\underline{F}_A(x)=\begin{cases}\frac12 x & \text{if } x\in\left[0,\frac23\right).\\ x & \text{otherwise.}\end{cases}$$

$$\overline{F}_A(x)=\begin{cases}x & \text{if } x\in\left[0,\frac13\right).\\ \frac12 x+\frac12 & \text{otherwise.}\end{cases}$$

Then, $P_{\Gamma_A}\left(\left(\frac13,\frac23\right]\right)=\frac13$. *However, it holds that:*

$$E_{\underline{F}_A,\overline{F}_A}\left(\left(\frac13,\frac23\right]\right)=\underline{F}_A\left(\frac23\right)-\overline{F}_A\left(\frac12\right)=\frac23-\frac23=0.$$

Furthermore:

$$E_{\underline{F}_A,\overline{F}_A}\left(\frac{1}{3},\frac{2}{3}\right) = \sup_{B\subseteq[\frac{1}{3},\frac{2}{3}],B\in H} E_{\underline{F}_A,\overline{F}_A}(B) \leq E_{\underline{F}_A,\overline{F}_A}\left(\frac{1}{3},\frac{2}{3}\right) = 0.$$

Thus, the natural extension is less informative than the original lower probability.

Next we show that the lower probability and the natural extension defined of the p-box coincide when $\mu_A = 0$.

Proposition 5.72 *Consider the initial space* $([0,1],\beta_{[0,1]},\lambda_{[0,1]})$ *and the random interval defined from an IF-set* A *with* $\mu_A = 0$. *Then,* $E_{\underline{F}_A,\overline{F}_A} = P_{\Gamma_A}$.

Proof We know that $\mu_A = 0$ implies that $\overline{F}_A = 1$. Let us prove the equality between the natural extension and the lower probability following several steps:

1. Let B be a set on H. We have several cases:
   - Assume that $B = [0,x]$. Then:
   $$P_{\Gamma_A}([0,x]) = P(\{\omega : \Gamma(\omega) \subseteq [0,x]\}) = \underline{F}_A(x).$$
   $$E_{\underline{F}_A,\overline{F}_A}([0,x]) = \underline{F}_A(x).$$
   - Assume now that $B = [0,x_1) \cup [x_2,x_3) \cup \dots \cup [x_{2k},x_{2k+1}]$, with $x_1 < x_2 < \dots < x_n$. Then:
   $$P_{\Gamma_A}(B) = P(\{\omega : \Gamma_A(\omega) \subseteq B\}) = P(\{\omega : \Gamma_A(\omega) \subseteq [0,x_1]\}) = \underline{F}_A(x_1).$$
   $$\begin{aligned} E_{\underline{F}_A,\overline{F}_A}(B) &= \underline{F}_A(x_1) + \sum_{i=1}^{n} \max\{0, \underline{F}_A(x_{2i+1}) - \overline{F}_A(x_{2i})\} \\ &= \underline{F}_A(x_1) + \sum_{i=1}^{n} \max\{0, \underline{F}_A(x_{2i+1}) - 1\} = \underline{F}_A(x_1). \end{aligned}$$
   - Finally, assume that $B = (x_1,x_2] \cup \dots \cup (x_{2n},x_{2n+1}]$, with $x_1 < x_2 < \dots < x_n$. Then:
   $$P_{\Gamma_A}(B) = P(\{\omega : \Gamma_A(\omega) = [0, 1-\nu_A(\omega)] \subseteq B\}) = 0.$$
   $$\begin{aligned} E_{\underline{F}_A,\overline{F}_A}(B) &= \sum_{i=1}^{n} \max\{0, \underline{F}_A(x_{2i+1}) - \overline{F}_A(x_{2i})\} \\ &= \sum_{i=1}^{n} \max\{0, \underline{F}_A(x_{2i+1}) - 1\} = 0. \end{aligned}$$

   Then, $E_{\underline{F}_A,\overline{F}_A}$ and P_{Γ_A} coincide for elements in H.

2. Consider $C \subseteq [0,1]$. Denote by $x = \sup\{x : [0,x] \subseteq C\}$. We have several cases:

- Assume that $\{x : [0, x] \quad C\} = \quad$, that means that $0/ \quad C$. Then, $0/ \quad B$ for every $B \quad H$, and then $E_{\underline{E}_A, \overline{F}_A}(B) = 0$. Thus, we conclude that

$$E_{\underline{E}_A, \overline{F}_A}(C) = \sup_{B \quad C, B \quad H} (B) = 0.$$

Furthermore, since $0/ \quad C$, $P_{\Gamma_A}(C) = 0$.

- Now, assume that $x \quad = \max\{x : [0, x] \quad C\}$, that means that $0 \quad C$ and there is x such that $[0, x \quad] \quad C$ but $[0, x \quad + \varepsilon] \quad C$ for any $\varepsilon > 0$. Then:

$$P_{\Gamma_A}(C) = P(\{\omega : \Gamma_A(\omega) \quad C\}) = P(\{\omega : \Gamma_A(\omega) \quad [0, x \quad]\}) = \overline{F}_A(x \quad).$$
$$E_{\underline{E}_A, \overline{F}_A}([0, x \quad]) = \overline{F}_A(x \quad).$$

Furthermore, as in the previous case:

$$E_{\underline{E}_A, \overline{F}_A}([0, x \quad]) = \overline{E}_{\underline{E}_A, \overline{F}_A}(B)$$

for any B such that $[0, x \quad] \quad B$, and consequently

$$E_{\underline{E}_A, \overline{F}_A}(C) = \overline{E}_{\underline{E}_A, \overline{F}_A}([0, x \quad]) = \overline{F}_A(x \quad).$$

- Finally, assume that x is a supremum, not a maximum, that is: $[0, x \quad) \quad C$ but $x \quad / \; C$. Then:

$$\begin{aligned} P_{\Gamma_A}(C) &= P(\{\omega : \Gamma_A(\omega) \quad [0, x \quad)\}) = \lim_{\varepsilon \to 0} P(\{\omega : 1 - \nu_A(\omega) \leq x \quad - \varepsilon\}) \\ &= \lim_{\varepsilon \to 0} \underline{E}_A(x \quad - \varepsilon) = \lim_{\varepsilon \to 0} P_{\Gamma_A}([0, x \quad - \varepsilon]) \\ &= \lim_{\varepsilon \to 0} E_{\underline{E}_A, \overline{F}_A}([0, x \quad - \varepsilon]) \\ &= \sup_{B \quad [0, x \quad), B \quad H} E_{\underline{E}_A, \overline{F}_A}(B) = \overline{E}_{\underline{E}_A, \overline{F}_A}([0, x \quad)). \end{aligned}$$

In addition, every $B \quad H$ such that $[0, x \quad) \quad B$ satisfies that $E_{\underline{E}_A, \overline{F}_A}(B) = E_{\underline{E}_A, \overline{F}_A}([0, x \quad))$. Then, the lower probability and the natural extension coincide. ∎

We could think that the lower probability and the natural extension of the associated p-box also coincide when the bounds of the random interval are strictly comonotone functions. However, we can find examples where such equality does not hold.

Example 5.73 *Consider the random interval* Γ_A *defined on* $([0, 1], \beta_{[0,1]}, \lambda_{[0,1]})$ *by:*

$$\Gamma_A(\omega) = \begin{cases} \left[\frac{1}{2} - \omega, 1 - \omega\right] & \text{if } \omega \quad \left[0, \frac{1}{4}\right). \\ \left[\frac{1}{4}, \frac{3}{4}\right] & \text{if } \omega \quad \left[\frac{1}{4}, \frac{3}{4}\right). \\ \left[\omega - \frac{1}{2}, \omega\right] & \text{if } \omega \quad \left[\frac{3}{4}, 1\right]. \end{cases}$$

Since $\mu_A(\omega) = (1-\nu_A(\omega)) - \frac{1}{2}$, *we see that* μ_A *and* $1-\nu_A$ *are strictly comonotone. Its associated p-box is defined by:*

$$\overline{F}_A(t) = \begin{cases} 0 & \text{if } t \in [0, \frac{1}{4}], \\ 2t & \text{if } t \in [\frac{1}{4}, \frac{1}{2}], \\ 1 & \text{if } t \in [\frac{1}{2}, 1], \end{cases} \quad \text{and } \underline{F}_A(t) = \begin{cases} 0 & \text{if } t \in [0, \frac{3}{4}]. \\ 2t-1 & \text{if } t \in [\frac{3}{4}, 1]. \end{cases}$$

Let us compute P_{Γ_A} *and* $E_{\underline{F}_A,\overline{F}_A}$ *for the set* $[\frac{1}{4}, \frac{7}{8}]$:

$$P_{\Gamma_A}\left([\tfrac{1}{4}, \tfrac{7}{8}]\right) = P\left(\omega : \Gamma_A(\omega) \subseteq [\tfrac{1}{4}, \tfrac{7}{8}]\right) = \tfrac{1}{4}.$$
$$E_{\underline{F}_A,\overline{F}_A}\left([\tfrac{1}{4}, \tfrac{7}{8}]\right) = \max\left(0, \underline{F}_A(\tfrac{1}{4}) - \overline{F}_A(\tfrac{7}{8})\right) = \tfrac{1}{4}.$$

Thus, they coincide. However, we are going to check that they do not agree on the set $(\frac{1}{4}, \frac{7}{8})$.

$$P_{\Gamma_A}\left((\tfrac{1}{4}, \tfrac{7}{8})\right) = P\left(\omega : \Gamma_A(\omega) \subseteq (\tfrac{1}{4}, \tfrac{7}{8})\right) = \frac{3}{4}.$$

By definition, $E_{\underline{F}_A,\overline{F}_A}\left((\frac{1}{4}, \frac{7}{8})\right) = \sup_{B \subseteq [\frac{1}{4}, \frac{7}{8}], B \in H} E_{\underline{F}_A,\overline{F}_A}(B)$. *But*

$$E_{\underline{F}_A,\overline{F}_A}(B) \leq E_{\underline{F}_A,\overline{F}_A}\left([\frac{1}{4}, \frac{7}{8}]\right) = \frac{1}{4}$$

for any $B \subseteq [\frac{1}{4}, \frac{7}{8}]$ *in* H. *Thus,*

$$P_{\Gamma_A}\left((\frac{1}{4}, \frac{7}{8})\right) > E_{\underline{F}_A,\overline{F}_A}\left((\frac{1}{4}, \frac{7}{8})\right).$$

5.2.2 Connection with other approaches

We now investigate the connection between the framework we have presented and other theories that can be found in the literature. For this aim, we first investigate the connection with the approach of Grzegorzewski and Mrowka ([86]) and then we establish a one-to-one relationship between IVF-sets, p-boxes and clouds.

Probabilities associated with IF-Sets

One of the most important works on the connection between IF-sets and imprecise probabilities is the work carried out in [86] on the probabilities of IF-sets. Given a probability space $(\Omega, \mathcal{A}, P)$, the probability associated with an IF-set A is a number of the interval

$$\left[\int_\Omega \mu_A \, dP, \int_\Omega 1-\nu_A \, dP\right]. \tag{5.10}$$

Using this definition, in [86] a link is established with probability theory by considering the appropriate operators in the spaces of real intervals and of intuitionistic fuzzy sets. Note that in this work it is assumed that we have a structure of probability space on Ω and that moreover the functions μ_A, ν_A are measurable, as we have done in Section 5.2.1.

Remark 5.74 *This definition generalises an earlier definition by Zadeh [215] for fuzzy events. He defined the probability of a fuzzy event μ_A by:*

$$P(\mu_A) = \int_\Omega \mu_A \, dP = E[\mu_A].$$

Although Zadeh proved that this definition satisfies the axioms of Kolmogorov when considering the minimum operator for making intersections, it was proved in [144] that this does not happen for any t-norm (see [100] for a complete review on t-norms). In fact, it was proved that every strict and continuous t-norm made Zadeh's probability to satisfy Kolmogorov axioms, while the Łukasiewicz operator is the only nilpotent and continuous t-norm that satisfies these axioms.

If we consider the random interval associated with the intuitionistic fuzzy set A in Equation (5.5), we can see that the interval in Equation (5.10) corresponds simply to the set of expectations of the measurable selections of Γ_A: it follows from [130, Theorem 14] that if we consider the mapping $id : [0,1] \to [0,1]$ then the *Aumann integral* [13] of $(id \circ \Gamma_A)$, defined on Equation (2.26), satisfies that

$$\left[\inf (A) \int (id \circ \Gamma_A) dP, \sup (A) \int (id \circ \Gamma_A) dP \right] = \left[(C) \int id \, dP_{\Gamma_A}, (C) \int id \, dP^{\Gamma_A} \right],$$

where (C) is used to denote the Choquet integral [39, 60] with respect to the non-additive measures $P_{\Gamma_A}, P^{\Gamma_A}$, respectively. Since on the other hand it is immediate to see that

$$\sup (A) \int (id \circ \Gamma_A) dP = \int (1 - \nu_A) dP$$

and

$$\inf (A) \int (id \circ \Gamma_A) dP = \int \mu_A \, dP,$$

we deduce that the probabilistic information about the intuitionistic fuzzy set A can be determined in particular by the lower and upper probabilities of its associated random interval. Note moreover that the Aumann integral of a random set is not convex in general, and it is only guaranteed to be so when the probability space $(\Omega, \mathcal{A}, P)$ is non-atomic.

A one-to-one relationship between p-boxes and IFS

In Section 5.2.1, we saw that the correspondence between interval-valued fuzzy sets and p-boxes on [0,1] is many-to-one, in the sense that many different IFS determine the same lower and upper distribution functions. In this section, we consider a subset of the class of IFS for which a bijection can be established with the set of p-boxes. In contradistinction to our work in Section 5.2.1, the p-box we shall establish here shall be established in the possibility space Ω, that we shall consider here to be the unit interval.

Denote by $IF([0,1])$ the set:

$$IF([0,1]) = \{A \quad IFSs(\Omega) \mid \mu_A \text{ increasing and } \nu_A \text{ decreasing}\}.$$

Denote also $F([0,1])$ the set of all p-boxes on [0,1], and let us define the correspondences:

$$\begin{array}{llcl} f_1: & F([0,1]) & \longrightarrow & IF([0,1]) \\ & (\underline{F},\overline{F}) & \to & A_{(\underline{F},\overline{F})} = (x, \underline{F}(x), 1-\overline{F}(x)) \end{array}$$

$$\begin{array}{llcl} f_2: & IF([0,1]) & \longrightarrow & F([0,1]) \\ & A & \to & (\mu_A, 1-\nu_A) \end{array}$$

We can see that every IFS A has an associated p-box $(\mu_A, 1-\nu_A)$. The interpretation here would be that $(\mu_A, 1-\nu_A)$ models the imprecise information about the distribution function associated with the set A, instead of about the membership function, as we did in Section 5.2.1.

The following properties follow immediately, and therefore their proof is omitted:

Proposition 5.75 *Let f_1, f_2 be the two correspondences between $F([0,1])$ and $IF([0,1])$ considered above. Then:*

(a) f_1, f_2 *are bijective, and* $f_1 = f_2^{-1}$.

(b) $f_1((\underline{F},\overline{F})) \quad \Gamma \quad \underline{F} = \overline{F}$.

(c) $f_2(A) = (F,F) \quad A \quad FS(\Omega)$.

Another property assures that there exists a relationship between application f_1 and the stochastic order $(FSD_{2,5})$:

$$(\underline{F}_1,\overline{F}_1) \quad {}_{FSD_{2,5}} \quad (\underline{F}_2,\overline{F}_2) \qquad f_1((\underline{F}_1,\overline{F}_1)) \qquad f_1((\underline{F}_2,\overline{F}_2)).$$

A one-to-one relationship between clouds and IFS

A similar correspondence can be made between intuitionistic fuzzy sets and clouds. Recall that a cloud is a pair of functions δ, π such that $\delta \leq \pi$ and there are $x, y \quad [0,1]$ such that $\delta(x) = 0$ and $\pi(y) = 1$. Let us denote by IF the following set:

$$IF = \{A \quad IFSs(\Omega) \mid \mu_A(x) = 0 \text{ and } \nu_A(y) = 0 \text{ for some } x, y \quad [0,1]\}.$$

Then, if we denote by $Cl([0,1])$ the set of all the clouds on [0,1], the following functions can be defined:

$$g_1 : Cl([0,1]) \longrightarrow IF([0,1])$$
$$(\delta, \pi) \rightarrow A_{(\delta,\pi)} = (x, \delta(x), 1 - \pi(x))$$

$$g_2 : IF([0,1]) \longrightarrow Cl([0,1])$$
$$A \rightarrow (\mu_A, 1 - \nu_A)$$

A cloud (δ, π) is called *thin* ([168]), when $\delta = \pi$; in that case, its associated IVF-sets by g_1 becomes $(x, \delta, 1 - \delta)$ $FS(\Omega)$, that is, a fuzzy set.

This is consistent in the sense that, given a possibility distribution π, it has an associated fuzzy set $\mu(x) := \pi(x)$. Thus, this is a more general approach that contains the relationship between fuzzy sets and possibility distribution as a particular case.

Another particular type of clouds are the *fuzzy clouds*, for which $\delta = 0$. In such a case the associated IFS is $(x, 0, 1 - \pi)$.

Some immediate properties of the above correspondences are the following:

Proposition 5.76 *Let g_1, g_2 be the correspondences between $Cl([0,1])$ and $IF([0,1])$ considered above. The following conditions hold:*

(a) $g_1((\delta, \pi))$ $FS(\Omega)$ $\delta = \pi$ (δ, π) *is a thin cloud.*

(b) $g_2(A) = (\delta, \delta)$ A $FS(\Omega)$.

(c) g_1, g_2 *are bijective, and* $g_1 = g_2^{-1}$.

The above correspondence is related to the connection between clouds and imprecise probabilities established in [65], where the credal set associated with a cloud (δ, π) is the set of probability measures on Ω satisfying $M((\pi, 1 - \delta)) = M(\pi) \cap M(1 - \delta)$, where $M(\pi)$ (respectively $M(1 - \delta)$) is the credal set associated with the possibility measure π (respectively $1 - \delta$).

5.3 Applications

In the previous sections we have presented a theoretical study of comparison measures for intuitionistic fuzzy sets, focusing in the study of IF-divergences, and we have also investigated the connection between IVF-sets and imprecise probabilities.

Now we shall present some possible applications of the theories we have developed. On one hand, we will see how IF-divergences can be applied in multiple attribute decision making ([211]), and we will outline some examples of application in pattern recognition ([92, 93, 114]). On the other hand, we shall see how the connection between IVF-sets and imprecise probabilities allows us to extend stochastic dominance to the comparison more than two sets of cumulative distribution functions.

5.3.1 Application to pattern recognition

One interesting area of application of comparison measures between IF-sets is in pattern recognition ([92, 93, 114]). Let us consider a universe $\Omega = \{\omega_1, \ldots, \omega_n\}$, and assume the patterns $A_1, \ldots, A_m$, that are represented by IF-sets. Then:

$$A_j = \{(\omega_i, \mu_{A_j}(\omega_i), \nu_{A_j}(\omega_i)) \mid i = 1, \ldots, n\}, \text{ for } j = 1, \ldots, m.$$

If B is a sample that is also represented by an IF-set, and we want to classify it into one of the patterns, we can measure the difference between B and A_i:

$$D_{\text{IFS}}(A_1, B), \ldots, D_{\text{IFS}}(A_m, B),$$

where D_{IFS} can be an IF-divergence or an IF-dissimilarity. Finally, we associate B to the pattern A_j whenever $D_{\text{IFS}}(A_j, B) = \min_{i=1,\ldots,m} (D_{\text{IFS}}(A_i, B))$, i.e., we classify B into the pattern from which it differs the least.

Example 5.77 ([114, Section 4]) *Consider a possibility space with three elements, $\Omega = \{\omega_1, \omega_2, \omega_3\}$, and the following three patterns:*

$$A_1 = \{(\omega_1, 0.1, 0.1), (\omega_2, 0.5, 0.4), (\omega_3, 0.1, 0.9)\}.$$
$$A_2 = \{(\omega_1, 0.5, 0.5), (\omega_2, 0.7, 0.3), (\omega_3, 0, 0.8)\}.$$
$$A_3 = \{(\omega_1, 0.7, 0.2), (\omega_2, 0.1, 0.8), (\omega_3, 0.4, 0.4)\}$$

Assume that a sample $B = \{(\omega_1, 0.4, 0.4), (\omega_2, 0.6, 0.2), (\omega_3, 0, 0.8)\}$ is given, and let us consider the Hamming and the Hausdorff distances for IF-sets. We obtain the following results.

$$l_{\text{IFS}}(A_1, B) = 1, \quad l_{\text{IFS}}(A_2, B) = 0.4, \quad l_{\text{IFS}}(A_3, B) = 1.3,$$
$$d_{\text{H}}(A_1, B) = 0.6, \quad d_{\text{H}}(A_2, B) = 0.2, \quad d_{\text{H}}(A_3, B) = 1.3.$$

Thus, both distances classify B into the pattern A_2, because

$$l_{\text{IFS}}(A_2, B) \le l_{\text{IFS}}(A_1, B), l_{\text{IFS}}(A_3, B).$$
$$d_{\text{H}}(A_2, B) \le d_{\text{H}}(A_1, B), d_{\text{H}}(A_3, B).$$

In the framework of pattern recognition it is usually assumed that every point u_i in the universe has the same weight, that is, $\alpha_i = \frac{1}{n}$ for $i = 1, \ldots, n$. However, it is possible that the weight vector $\alpha = (\alpha_1, \ldots, \alpha_n)$ is not constant, that is, $\alpha_i \ge 0$ for $i = 1, \ldots, n$ and $\alpha_1 + \ldots + \alpha_n = 1$.

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6
$\mu_{C_1}(\omega_i)$	0.739	0.033	0.188	0.492	0.020	0.739
$\nu_{C_1}(\omega_i)$	0.125	0.818	0.626	0.358	0.628	0.125
$\mu_{C_2}(\omega_i)$	0.124	0.030	0.048	0.136	0.019	0.393
$\nu_{C_2}(\omega_i)$	0.665	0.825	0.800	0.648	0.823	0.653
$\mu_{C_3}(\omega_i)$	0.449	0.662	1.000	1.000	1.000	1.000
$\nu_{C_3}(\omega_i)$	0.387	0.298	0.000	0.000	0.000	0.000
$\mu_{C_4}(\omega_i)$	0.280	0.521	0.470	0.295	0.188	0.735
$\nu_{C_4}(\omega_i)$	0.715	0.368	0.423	0.658	0.806	0.118
$\mu_{C_5}(\omega_i)$	0.326	1.000	0.182	0.156	0.049	0.675
$\nu_{C_5}(\omega_i)$	0.452	0.000	0.725	0.765	0.896	0.263
$\mu_B(\omega_i)$	0.629	0.524	0.210	0.218	0.069	0.658
$\nu_B(\omega_i)$	0.303	0.356	0.689	0.753	0.876	0.256

Table 5.2: Six kinds of materials are represented by IF-sets.

To deal with this situation, we propose the following method. Let us consider a local IF-divergence D_{IFS}, and for every point u_i let us compute the following:

$$D_{IFS}(A_j, B) - D_{IFS}(A_j \cup \{\omega_i\}, B \cup \{\omega_i\}) = h_{IFS}(\mu_{A_j}(\omega_i), \nu_{A_j}(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)).$$

Then, for every $j \in \{1, \ldots, m\}$ we have that

$$\begin{aligned} d(A_j, B) &= \sum_{i=1}^{n} \omega_i \left(D_{IFS}(A_j, B) - D_{IFS}(A_j \cup \{\omega_i\}, B \cup \{\omega_i\})\right) \\ &= \sum_{i=1}^{n} \alpha_i h_{IFS}(\mu_{A_j}(\omega_i), \nu_{A_j}(\omega_i), \mu_B(\omega_i), \nu_B(\omega_i)). \end{aligned}$$

Then, we classify the sample B into the pattern A_j if

$$d(A_j, B) = \min_{i=1,\ldots,m} (d(A_i, B)).$$

Example 5.78 ([206, Example 4.2]) *Consider five kinds of mineral fields, each of them featured by the content of six minerals and containing one kind of typical hybrid mineral. The five kinds of typical hybrid mineral are represented by IF-sets C_1, C_2, C_3, C_4 and C_5 in $\Omega = \{\omega_1, \ldots, \omega_6\}$, respectively. Assume that we are given another kind of hybrid mineral B, and that we want to classify it into one of the aforementioned mineral fields. Assume that the IF-sets C_i and B are defined in Table 5.2, and that our experts have established the following weight vector on Ω: $\alpha = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}\right)$. Let us use our method to classify B. If we consider the Hamming distance for IF-sets as local*

IF-divergence, we obtain that:

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6
$I_{IFS}(C_1, B) - I_{IFS}(C_1 \ \{\omega\}, B \ \{\omega\})$	0.178	0.491	0.085	0.395	0.297	0.131
$I_{IFS}(C_2, B) - I_{IFS}(C_2 \ \{\omega\}, B \ \{\omega\})$	0.505	0.494	0.162	0.187	0.103	0.397
$I_{IFS}(C_3, B) - I_{IFS}(C_3 \ \{\omega\}, B \ \{\omega\})$	0.180	0.138	0.790	0.782	0.931	0.342
$I_{IFS}(C_4, B) - I_{IFS}(C_4 \ \{\omega\}, B \ \{\omega\})$	0.412	0.012	0.266	0.095	0.119	0.138
$I_{IFS}(C_5, B) - I_{IFS}(C_5 \ \{\omega\}, B \ \{\omega\})$	0.303	0.476	0.036	0.062	0.020	0.024

whence

$$d(C_1, B) = \tfrac{1}{4}0.178 + \tfrac{1}{4}0.491 + \tfrac{1}{8}0.085 + \tfrac{1}{8}0.395 + \tfrac{1}{8}0.297 + \tfrac{1}{8}0.131 = 0.2808.$$

$$d(C_2, B) = \tfrac{1}{4}0.505 + \tfrac{1}{4}0.494 + \tfrac{1}{8}0.162 + \tfrac{1}{8}0.187 + \tfrac{1}{8}0.103 + \tfrac{1}{8}0.397 = 0.3559.$$

$$d(C_3, B) = \tfrac{1}{4}0.180 + \tfrac{1}{4}0.138 + \tfrac{1}{8}0.790 + \tfrac{1}{8}0.782 + \tfrac{1}{8}0.931 + \tfrac{1}{8}0.342 = 0.4351.$$

$$d(C_4, B) = \tfrac{1}{4}0.412 + \tfrac{1}{4}0.012 + \tfrac{1}{8}0.266 + \tfrac{1}{8}0.095 + \tfrac{1}{8}0.119 + \tfrac{1}{8}0.138 = 0.1833.$$

$$d(C_5, B) = \tfrac{1}{4}0.303 + \tfrac{1}{4}0.476 + \tfrac{1}{8}0.036 + \tfrac{1}{8}0.062 + \tfrac{1}{8}0.020 + \tfrac{1}{8}0.024 = 0.2125.$$

Thus, we classify B into the hybrid mineral C_4.

If we repeat the process with local IF-divergence d_H, we obtain the following:

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6
$d_H(C_1, B) - d_H(C_1 \ \{\omega\}, B \ \{\omega\})$	0.178	0.491	0.063	0.395	0.248	0.131
$d_H(C_2, B) - d_H(C_2 \ \{\omega\}, B \ \{\omega\})$	0.505	0.494	0.162	0.105	0.053	0.397
$d_H(C_3, B) - d_H(C_3 \ \{\omega\}, B \ \{\omega\})$	0.180	0.138	0.790	0.782	0.931	0.342
$d_H(C_4, B) - d_H(C_4 \ \{\omega\}, B \ \{\omega\})$	0.412	0.012	0.266	0.095	0.119	0.138
$d_H(C_5, B) - d_H(C_5 \ \{\omega\}, B \ \{\omega\})$	0.303	0.476	0.036	0.062	0.020	0.017

Then:

$$d(C_1, B) = \tfrac{1}{4}0.178 + \tfrac{1}{4}0.491 + \tfrac{1}{8}0.063 + \tfrac{1}{8}0.395 + \tfrac{1}{8}0.248 + \tfrac{1}{8}0.131 = 0.2719.$$

$$d(C_2, B) = \tfrac{1}{4}0.505 + \tfrac{1}{4}0.494 + \tfrac{1}{8}0.162 + \tfrac{1}{8}0.105 + \tfrac{1}{8}0.053 + \tfrac{1}{8}0.397 = 0.3394.$$

$$d(C_3, B) = \tfrac{1}{4}0.180 + \tfrac{1}{4}0.138 + \tfrac{1}{8}0.790 + \tfrac{1}{8}0.782 + \tfrac{1}{8}0.931 + \tfrac{1}{8}0.342 = 0.4351.$$

$$d(C_4, B) = \tfrac{1}{4}0.412 + \tfrac{1}{4}0.012 + \tfrac{1}{8}0.266 + \tfrac{1}{8}0.095 + \tfrac{1}{8}0.119 + \tfrac{1}{8}0.138 = 0.1833.$$

$$d(C_5, B) = \tfrac{1}{4}0.303 + \tfrac{1}{4}0.476 + \tfrac{1}{8}0.036 + \tfrac{1}{8}0.062 + \tfrac{1}{8}0.020 + \tfrac{1}{8}0.017 = 0.2116,$$

and we conclude that we also should classify B into the hybrid mineral C_4.

5.3.2 Application to decision making

In [211], Xu showed how measures of similarity for IF-sets (and, consequently, also IF-dissimilarities) can be applied within multiple attribute decision making. Let us overview the main aspects of this application.

We use the following notation: let $A = \{A_1, \ldots, A_m\}$ denote a set of m alternatives, let $C = \{C_1, \ldots, C_n\}$ be a set of attributes and let $\alpha = \{\alpha_1, \ldots, \alpha_n\}$ be its associated weight vector (i.e., it holds that $\alpha_i \geq 0$ for every $i = 1, \ldots, n$ and that $\alpha_1 + \ldots + \alpha_n = 1$).

Every alternative A_i can be represented by means of an IF-set:

$$A_i = \{(C_j, \mu_{A_i}(C_j), \nu_{A_i}(C_j)) \mid j = 1, \ldots, n\}.$$

Thus, $\mu_{A_i}(C_j)$ and $\nu_{A_i}(C_j)$ stand for the degree in which alternative A_i agrees and does not agree with characteristic C_j, respectively.

Xu ([211]) defined the IF-sets A^+ and A^- in the following way:

$$\begin{aligned} A^+ &= \{(C_j, \mu_{A^+}(C_j), \nu_{A^+}(C_j)) \mid j = 1, \ldots, n\} \text{ and} \\ A^- &= \{(C_j, \mu_{A^-}(C_j), \nu_{A^-}(C_j)) \mid j = 1, \ldots, n\}, \end{aligned}$$

where

$$\mu_{A^+}(C_j) = \max_{i=1,\ldots,m} (\mu_{A_i}(C_j)), \quad \nu_{A^+}(C_j) = \min_{i=1,\ldots,m} (\nu_{A_i}(C_j)), \tag{5.11}$$

$$\mu_{A^-}(C_j) = \min_{i=1,\ldots,m} (\mu_{A_i}(C_j)), \quad \nu_{A^-}(C_j) = \max_{i=1,\ldots,m} (\nu_{A_i}(C_j)), \tag{5.12}$$

that is, $A^+ = \overset{m}{\underset{i=1}{}} A_i$ and $A^- = \overset{m}{\underset{i=1}{}} A_i$.

These IF-sets can be interpreted as the "optimal" and the "least optimal" alternatives. Therefore, the preferred alternative in A would the one that is simultaneously more similar to A^+ and more different to A^-.

In order to measure how different is A_i to both A^+ and A^-, Xu considered some different functions, such as:

$$D(A^+, A_i) = \sum_{j=1}^{n} \alpha_j \Big(|\mu_{A^+}(C_j) - \mu_{A_i}(C_j)|^\beta + |\nu_{A^+}(C_j) - \nu_{A_i}(C_j)|^\beta + |\pi_{A^+}(C_j) - \pi_{A_i}(C_j)|^\beta \Big)^{\frac{1}{\beta}}$$

and

$$D(A^-, A_i) = \sum_{j=1}^{n} \alpha_j \Big(|\mu_{A^-}(C_j) - \mu_{A_i}(C_j)|^\beta + |\nu_{A^-}(C_j) - \nu_{A_i}(C_j)|^\beta + |\pi_{A^+}(C_j) - \pi_{A_i}(C_j)|^\beta \Big)^{\frac{1}{\beta}}.$$

Besides, Xu consider the quotient:

$$d_i = \frac{D(A^+, A_i)}{D(A^+, A_i) + D(A^-, A_i)}.$$

Then, the greater the value d_i, the better the alternative A_i.

Next we propose a modification of the above method. Let us consider a local IF-divergence D_{IFS}, so that for every pair of IF-sets A and B, $D_{IFS}(A, B)$ can be expressed by:

$$D_{IFS}(A, B) = \sum_{i=1}^{n} h_{IFS}(\mu_A(C_i), \nu_A(C_i), \mu_B(C_i), \nu_B(C_i)).$$

We consider the IF-set A_i, that represents the i-th alternative, and for every $j \in \{1, \ldots, n\}$ we compute the following:

$$D_{IFS}(A^+, A_i) - D_{IFS}(A^+ \setminus \{C_j\}, A_i \setminus \{C_j\}) = h_{IFS}(\mu_{A^+}(C_j), \nu_{A^+}(C_j), \mu_{A_i}(C_j), \nu_{A_i}(C_j)).$$

This quantity measures how different A^+ and A_i are with respect to element C_j. Then, we can compute the difference between A_i and A^+:

$$d(A_i, A^+) = \sum_{j=1}^{n} \alpha_j h_{IFS}(\mu_{A^+}(C_j), \nu_{A^+}(C_j), \mu_{A_i}(C_j), \nu_{A_i}(C_j)).$$

In this way $d(A_i, A^+)$ measures how much difference there is between A_i and the optimal set A^+.

Similarly, we can compute the difference between A_i and A^-:

$$d(A_i, A^-) = \sum_{j=1}^{n} \alpha_j h_{IFS}(\mu_{A^-}(C_j), \nu_{A^-}(C_j), \mu_{A_i}(C_j), \nu_{A_i}(C_j)).$$

Thus, $d(A_i, A^-)$ measures how much different is A_i from the least optimal A^-.

Therefore, if we consider a map $f : [0, \infty) \times [0, \infty) \to [0, \infty)$ that is decreasing in the first component and increasing on the second one, we obtain the following value a_i for alternative A_i:

$$a_i = f(d(A_i, A^+), d(A_i, A^-)).$$

Thus, the greater the value of a_i, the more preferred is the alternative A_i.

We can see that we can choose the function f depending on the part we are more interested in: the difference between A_i and the optimum A^+ or the difference between A_i and the least optimum A^-. The following examples illustrate this fact.

Example 5.79 ([211, Section 4]) *A city is planning to build a library, and the city commissioner has to determine the air-conditioning system to be installed in the library. The builder offers the commissioner five feasible alternatives A_i, which might be adapted to the physical structure of the library. Suppose that three attributes C_1 (economic), C_2 (functional) and C_3 (operational) are taken into consideration in the installation problem,* *and that the weight vector of the attributes* C_j *is* $\alpha = (0.3, 0.5, 0.2)$. *Assume moreover that the characteristics of the alternatives* A_i *are represented by the following IF-sets:*

$$A_1 = \{(C_1, 0.2, 0.4), (C_2, 0.7, 0.1), (C_3, 0.6, 0.3)\}$$
$$A_2 = \{(C_1, 0.4, 0.2), (C_2, 0.5, 0.2), (C_3, 0.8, 0.1)\}$$
$$A_3 = \{(C_1, 0.5, 0.4), (C_2, 0.6, 0.2), (C_3, 0.9, 0)\}$$
$$A_4 = \{(C_1, 0.3, 0.5), (C_2, 0.8, 0.1), (C_3, 0.7, 0.2)\}$$
$$A_5 = \{(C_1, 0.8, 0.2), (C_2, 0.7, 0), (C_3, 0.1, 0.6)\}$$

For these IF-sets, the corresponding A^+ *and* A^- *are given by:*

$$A^+ = \{(C_1, 0.8, 0.2), (C_2, 0.8, 0), (C_3, 0.9, 0)\}$$
$$A^- = \{(C_1, 0.2, 0.5), (C_2, 0.5, 0.2), (C_3, 0.1, 0.6)\}$$

Then, if we consider the Hamming distance for IF-sets (see Subsection 5.1.3), we obtain the following:

	C_1	C_2	C_3
$l_{\mathrm{IFS}}(A_1, A^+) - l_{\mathrm{IFS}}(A_1\ \{C_j\}, A^+\ \{C_j\})$	1.2	0.2	0.6
$l_{\mathrm{IFS}}(A_1, A^-) - l_{\mathrm{IFS}}(A_1\ \{C_j\}, A^-\ \{C_j\})$	0.2	0.4	1
$l_{\mathrm{IFS}}(A_2, A^+) - l_{\mathrm{IFS}}(A_2\ \{C_j\}, A^+\ \{C_j\})$	0.8	0.6	0.2
$l_{\mathrm{IFS}}(A_2, A^+) - l_{\mathrm{IFS}}(A_2\ \{C_j\}, A^+\ \{C_j\})$	0.6	0	1.4
$l_{\mathrm{IFS}}(A_3, A^+) - l_{\mathrm{IFS}}(A_3\ \{C_j\}, A^+\ \{C_j\})$	0.6	0.4	0
$l_{\mathrm{IFS}}(A_3, A^+) - l_{\mathrm{IFS}}(A_3\ \{C_j\}, A^+\ \{C_j\})$	0.6	0.2	1.6
$l_{\mathrm{IFS}}(A_4, A^+) - l_{\mathrm{IFS}}(A_4\ \{C_j\}, A^+\ \{C_j\})$	1	0.2	0.4
$l_{\mathrm{IFS}}(A_4, A^+) - l_{\mathrm{IFS}}(A_4\ \{C_j\}, A^+\ \{C_j\})$	0.2	0.6	1.2
$l_{\mathrm{IFS}}(A_5, A^+) - l_{\mathrm{IFS}}(A_5\ \{C_j\}, A^+\ \{C_j\})$	0	0.2	1.6
$l_{\mathrm{IFS}}(A_5, A^+) - l_{\mathrm{IFS}}(A_5\ \{C_j\}, A^+\ \{C_j\})$	1.2	0.4	0

Thus:

$$d(A_1, A^+) = 0.3\ 1.2 + 0.5\ 0.2 + 0.2\ 0.6 = 0.58.$$
$$d(A_1, A^-) = 0.3\ 0.2 + 0.5\ 0.4 + 0.2\ 1 = 0.46.$$
$$d(A_2, A^+) = 0.3\ 0.8 + 0.5\ 0.6 + 0.2\ 0.2 = 0.58.$$
$$d(A_2, A^-) = 0.3\ 0.6 + 0.5\ 0 + 0.2\ 1.4 = 0.46.$$
$$d(A_3, A^+) = 0.3\ 0.6 + 0.5\ 0.4 + 0.2\ 0 = 0.38.$$
$$d(A_3, A^-) = 0.3\ 0.6 + 0.5\ 0.2 + 0.2\ 1.6 = 0.6.$$
$$d(A_4, A^+) = 0.3\ 1 + 0.5\ 0.2 + 0.2\ 0.4 = 0.48.$$
$$d(A_4, A^-) = 0.3\ 0.2 + 0.5\ 0.6 + 0.2\ 1.2 = 0.6.$$
$$d(A_5, A^+) = 0.3\ 0 + 0.5\ 0.2 + 0.2\ 1.6 = 0.42.$$
$$d(A_5, A^-) = 0.3\ 1.2 + 0.5\ 0.4 + 0.2\ 0 = 0.56.$$

Assume that we want to choose the alternative that is, at the same time, more similar to A^+ *and less similar to the worst case* A^-. *In such a case we can consider the function* f *given by* $f(x, y) = \frac{1}{2}\left(\frac{1}{x} + y\right)$. *We can see that this function take into account the difference* *between A_i and A^+ and between A_i and A^-. We obtain the fol lowing results:*

$$a_1 = f(d(A_1,A^+), d(A_1,A^-)) = \frac{1}{2}\,\frac{1}{0.58} + 0.46 = 1.09.$$
$$a_2 = f(d(A_2,A^+), d(A_2,A^-)) = \frac{1}{2}\,\frac{1}{0.58} + 0.46 = 1.09.$$
$$a_3 = f(d(A_3,A^+), d(A_3,A^-)) = \frac{1}{2}\,\frac{1}{0.38} + 0.6 = 1.62.$$
$$a_4 = f(d(A_4,A^+), d(A_4,A^+)) = \frac{1}{2}\,\frac{1}{0.48} + 0.6 = 1.34.$$
$$a_5 = f(d(A_5,A^+), d(A_5,A^+)) = \frac{1}{2}\,\frac{1}{0.42} + 0.56 = 1.47.$$

Assume next that we decide to choose the alternative that is more similar to the optimum A^+, regard less the difference from A^-. In that case, we may consider $f(x,y) = \frac{1}{x}$. This function only depends in the difference between A_i and the optimum A^+. We obtain the fol lowing result:

$$a_1 = f(d(A_1,A^+), d(A_1,A^-)) = \frac{1}{d(A_1,A^+)} = \frac{1}{0.58}.$$
$$a_2 = f(d(A_2,A^+), d(A_2,A^-)) = \frac{1}{d(A_2,A^+)} = \frac{1}{0.58}.$$
$$a_3 = f(d(A_3,A^+), d(A_3,A^-)) = \frac{1}{d(A_3,A^+)} = \frac{1}{0.38}.$$
$$a_4 = f(d(A_4,A^+), d(A_4,A^+)) = \frac{1}{d(A_4,A^+)} = \frac{1}{0.48}.$$
$$a_5 = f(d(A_5,A^+), d(A_5,A^+)) = \frac{1}{d(A_5,A^+)} = \frac{1}{0.42}.$$

Thus, A_3 A_5 A_4 A_1 A_2, and as a consequence the best alternative is A_3.

Final ly, assume we are interested in the alternative that differs more from the worst alternative A^-. In such a situation we should consider $f(x,y) = y$. This function only depends on the difference between A_i and A^-. We obtain the fol lowing results:

$$a_1 = f(d(A_1,A^+), d(A_1,A^-)) = d(A_1,A^-) = 0.46.$$
$$a_2 = f(d(A_2,A^+), d(A_2,A^-)) = d(A_2,A^-) = 0.46.$$
$$a_3 = f(d(A_3,A^+), d(A_3,A^-)) = d(A_3,A^-) = 0.6.$$
$$a_4 = f(d(A_4,A^+), d(A_4,A^+)) = d(A_4,A^-) = 0.6.$$
$$a_5 = f(d(A_5,A^+), d(A_5,A^+)) = d(A_5,A^-) = 0.56.$$

Thus, A_3 A_4 A_5 A_1 A_2. We conclude that in this case A_3 and A_4 are the preferred alternatives.

Example 5.80 *Consider the previous example, but now with the Hausdorff distance for IF-sets (see Section 5.1.3). Using the same IVF-sets, we obtain that:*

	C_1	C_2	C_3
$d_H(A_1,A^+) - d_H(A_1\ \{C_j\},A^+\ \{C_j\})$	0.6	0.1	0.3
$d_H(A_1,A^-) - d_H(A_1\ \{C_j\},A^-\ \{C_j\})$	0.3	0.2	0.5
$d_H(A_2,A^+) - d_H(A_2\ \{C_j\},A^+\ \{C_j\})$	0.4	0.3	0.1
$d_H(A_2,A^-) - d_H(A_2\ \{C_j\},A^+\ \{C_j\})$	0.3	0	0.7
$d_H(A_3,A^+) - d_H(A_3\ \{C_j\},A^+\ \{C_j\})$	0.3	0.2	0
$d_H(A_3,A^-) - d_H(A_3\ \{C_j\},A^+\ \{C_j\})$	0.3	0.1	0.8
$d_H(A_4,A^+) - d_H(A_4\ \{C_j\},A^+\ \{C_j\})$	0.5	0.1	0.2
$d_H(A_4,A^-) - d_H(A_4\ \{C_j\},A^+\ \{C_j\})$	0.3	0.3	0.6
$d_H(A_5,A^+) - d_H(A_5\ \{C_j\},A^+\ \{C_j\})$	0	0.1	0.8
$d_H(A_5,A^-) - d_H(A_5\ \{C_j\},A^+\ \{C_j\})$	0.6	0.2	0

Then:

$$d(A_1,A^+) = 0.3\ \ 0.6 + 0.5\ 0.1 + 0.2\ \ 0.3 = 0.29.$$
$$d(A_1,A^-) = 0.3\ \ 0.3 + 0.5\ \ 0.2 + 0.3\ 0.5 = 0.34.$$
$$d(A_2,A^+) = 0.3\ \ 0.4 + 0.5\ 0.3 + 0.3\ 0.1 = 0.3.$$
$$d(A_2,A^-) = 0.3\ \ 0.3 + 0.5\ \ 0 + 0.3\ \ 0.7 = 0.3.$$
$$d(A_3,A^+) = 0.3\ \ 0.3 + 0.5\ 0.2 + 0.3\ 0 = 0.19.$$
$$d(A_3,A^-) = 0.3\ \ 0.3 + 0.5\ 0.1 + 0.3\ \ 0.8 = 0.38.$$
$$d(A_4,A^+) = 0.3\ \ 0.5 + 0.5\ 0.1 + 0.3\ 0.2 = 0.26.$$
$$d(A_4,A^-) = 0.3\ \ 0.3 + 0.5\ 0.3 + 0.3\ \ 0.6 = 0.42.$$
$$d(A_5,A^+) = 0.3\ \ 0 + 0.5\ \ 0.1 + 0.3\ 0.8 = 0.29.$$
$$d(A_5,A^-) = 0.3\ \ 0.6 + 0.5\ 0.2 + 0.3\ \ 0 = 0.28.$$

As before, we first look for the alternative that is, at the same time, more similar to the optimum A^+ andless similarto theleast optimum A^-. For thisaim wecan consider the function $f(x,y) = \frac{1}{2}\ \frac{1}{x} + y$ *. It holdsthat:*

$$a_1 = f(d(A_1,A^+), d(A_1,A^-)) = \frac{1}{2}\ \frac{1}{0.29} + 0.34 = 3.79.$$
$$a_2 = f(d(A_2,A^+), d(A_2,A^-)) = \frac{1}{2}\ \frac{1}{0.3} + 0.3 = 3.63.$$
$$a_3 = f(d(A_3,A^+), d(A_3,A^-)) = \frac{1}{2}\ \frac{1}{0.19} + 0.38 = 5.64.$$
$$a_4 = f(d(A_4,A^+), d(A_4,A^+)) = \frac{1}{2}\ \frac{1}{0.26} + 0.42 = 4.27.$$
$$a_5 = f(d(A_5,A^+), d(A_5,A^+)) = \frac{1}{2}\ \frac{1}{0.29} + 0.28 = 3.72.$$

Then $A_3\ \ A_4\ \ A_1\ \ A_5\ \ A_2$, and therefore A_3 is thepreferred alternative.

Next, we seek for the alternative that is more similar to the optimal A^+. Apossible *function f for this scenario is $f(x,y) = \frac{1}{x}$. In such a case:*

$$a_1 = f(d(A_1, A^+), d(A_1, A^-)) = \frac{1}{d(A_1, A^+)} = \frac{1}{0.29}.$$
$$a_2 = f(d(A_2, A^+), d(A_2, A^-)) = \frac{1}{d(A_2, A^+)} = \frac{1}{0.3}.$$
$$a_3 = f(d(A_3, A^+), d(A_3, A^-)) = \frac{1}{d(A_3, A^+)} = \frac{1}{0.19}.$$
$$a_4 = f(d(A_4, A^+), d(A_4, A^+)) = \frac{1}{d(A_4, A^+)} = \frac{1}{0.26}.$$
$$a_5 = f(d(A_5, A^+), d(A_5, A^+)) = \frac{1}{d(A_5, A^+)} = \frac{1}{0.29}.$$

Then, it holds that A_3 A_4 A_1 A_5 A_2, and therefore alternative A_3 is the preferred one.

Finally, if we look for the alternative that differs more from the worst possibility A^-, we can choose $f(x,y) = y$. In that case,

$$a_1 = f(d(A_1, A^+), d(A_1, A^-)) = d(A_1, A^-) = 0.34.$$
$$a_2 = f(d(A_2, A^+), d(A_2, A^-)) = d(A_2, A^-) = 0.3.$$
$$a_3 = f(d(A_3, A^+), d(A_3, A^-)) = d(A_3, A^-) = 0.38.$$
$$a_4 = f(d(A_4, A^+), d(A_4, A^+)) = d(A_4, A^-) = 0.42.$$
$$a_5 = f(d(A_5, A^+), d(A_5, A^+)) = d(A_5, A^-) = 0.28.$$

We conclude that A_4 A_3 A_1 A_2 A_5, whence A_4 is the best alternative.

5.3.3 Using IF-divergences to extend stochastic dominance

Consider now the problem of comparing more than two random variables. In Section 3.3 we mentioned that both stochastic dominance and statistical preference are methods for the pairwise comparison of random variables, and we proposed a generalization of statistical preference for comparing more than two random variables, based on an extension of the probabilistic relation defined in Equation (2.7). Now, based on the IF-divergences and due to the connection between IF-sets and imprecise probabilities we have investigated in Section 5.2, we propose a method that allows us to compare n p-boxes in order to obtain an order between them.

In order to do this, consider n p-boxes $(\underline{F}_1, \overline{F}_1), \ldots, (\underline{F}_n, \overline{F}_n)$. For each p-box $(\underline{F}_i, \overline{F}_i)$, define the random interval Γ_i by $\Gamma_i(\omega) = [U_i(\omega), V_i(\omega)]$ where U_i and V_i are the quantile functions of $\underline{F}_i$ and $\overline{F}_i$, respectively. Then, for each p-box $(\underline{F}_i, \overline{F}_i)$ we have an associated random interval that we can understand as a random interval defined from an IF-set A_i. Thus, we can apply the method described in Section 5.3.2 to obtain the p-box closer to the "optimal" p-box, that is the one associated with A^+, and more distant to the "less optimal" p-box, that is the one associated with A^-.

Remark 5.81 *During this section we have investigated measures of comparison defined on finite spaces, according to the usual framework. However, all the measures we have studied can be extended to any space, non-necessarily finite. For instance, when dealing with local IF-divergences, they could be defined from* $[a, b]$ *to* $\mathbb{R}$ *by using the Lebesgue measure* $\lambda_{[a,b]}$ *in* $[a, b]$:

$$D_{\text{IFS}}(A, B) = \int_{[a,b]} h_{IFS}(\mu_A(\omega), \nu_A(\omega), \mu_B(\omega), \nu_B(\omega)) d\lambda_{[a,b]}.$$

In order to illustrate this method, we propose a numerical example based on the comparison of sets of Lorenz Curves as we made in Section 4.4.1.

Numerical example: comparison of Lorenz curves

In Section 4.4.1 we considered the Lorenz curves associated with several countries. Such data was illustrated in Table 4.2, and Table 4.3 showed the cumulative distribution functions associated with each Lorenz curve. Recall that we grouped the countries by continents/regions in the following way:

- Group 1: China, Japan, India.
- Group 2: Finland, Norway, Sweden.
- Group 3: Canada, USA.
- Group 4: FYR Macedonia, Greece.
- Group 5: Australia, Maldives.

Next table shows the p-boxes associated with these groups.

Group		**F(0.2)**	**F(0.4)**	**F(0.6)**	**F(0.8)**	**F(1)**
Group-1	$\overline{F}_1$	47.81	69.81	84.47	94.27	100
	$\underline{F}_1$	35.65	57.63	75.21	89.42	100
Group-2	$\overline{F}_2$	37.23	59.33	76.9	90.88	100
	$\underline{F}_2$	36.63	58.84	76.31	90.38	100
Group-3	$\overline{F}_3$	45.82	68.22	83.88	94.56	100
	$\underline{F}_3$	39.94	62.89	80.07	92.8	100
Group-4	$\overline{F}_4$	41.49	64.53	81.37	93.26	100
	$\underline{F}_4$	37.43	60.04	77.53	90.98	100
Group-5	$\overline{F}_5$	49.24	66.9	82.61	94.1	100
	$\underline{F}_5$	41.32	64.89	82.09	93.49	100

Assume now that we are interested in comparing all the groups of countries together. Then, following the steps of Section 5.3.2, denote by A_i the IF-set defined by $\mu_{A_i} = \overline{F}_i^{-1]}$ and $1 - \nu_{A_i} = \underline{F}_i^{-1]}$, that is, the IF-set defined by the quantile functions of $(\underline{F}_i, \overline{F}_i)$. These IF-sets are given by:

$$\mu_{A_1}(t) = \begin{cases} 0 & \text{if } t = 0. \\ 0.2 & \text{if } t \in (0, 47.81]. \\ 0.4 & \text{if } t \in (47.81, 69.81]. \\ 0.6 & \text{if } t \in (69.81, 84.47]. \\ 0.8 & \text{if } t \in (84.47, 94.27]. \\ 1 & \text{if } t \in (94.27, 100]. \end{cases} \qquad 1 - \nu_{A_1}(t) = \begin{cases} 0 & \text{if } t = 0. \\ 0.2 & \text{if } t \in (0, 35.64]. \\ 0.4 & \text{if } t \in (35.64, 57.63]. \\ 0.6 & \text{if } t \in (57.63, 75.21]. \\ 0.8 & \text{if } t \in (75.21, 89.42]. \\ 1 & \text{if } t \in (89.42, 100]. \end{cases}$$

$$\mu_{A_2}(t) = \begin{cases} 0 & \text{if } t = 0. \\ 0.2 & \text{if } t \in (0, 37.23]. \\ 0.4 & \text{if } t \in (37.23, 59.33]. \\ 0.6 & \text{if } t \in (59.33, 76.9]. \\ 0.8 & \text{if } t \in (76.9, 90.88]. \\ 1 & \text{if } t \in (90.88, 100]. \end{cases} \qquad 1 - \nu_{A_2}(t) = \begin{cases} 0 & \text{if } t = 0. \\ 0.2 & \text{if } t \in (0, 36.63]. \\ 0.4 & \text{if } t \in (36.63, 58.84]. \\ 0.6 & \text{if } t \in (58.84, 76.31]. \\ 0.8 & \text{if } t \in (76.31, 90.38]. \\ 1 & \text{if } t \in (90.38, 100]. \end{cases}$$

$$\mu_{A_3}(t) = \begin{cases} 0 & \text{if } t = 0. \\ 0.2 & \text{if } t \in (0, 45.82]. \\ 0.4 & \text{if } t \in (45.82, 68.22]. \\ 0.6 & \text{if } t \in (68.22, 83.88]. \\ 0.8 & \text{if } t \in (83.88, 94.56]. \\ 1 & \text{if } t \in (94.56, 100]. \end{cases} \qquad 1 - \nu_{A_3}(t) = \begin{cases} 0 & \text{if } t = 0. \\ 0.2 & \text{if } t \in (0, 39.94]. \\ 0.4 & \text{if } t \in (39.94, 62.89]. \\ 0.6 & \text{if } t \in (62.89, 80.07]. \\ 0.8 & \text{if } t \in (80.07, 92.8]. \\ 1 & \text{if } t \in (92.8, 100]. \end{cases}$$

$$\mu_{A_4}(t)=\begin{cases}0 & \text{if } t=0.\\ 0.2 & \text{if } t\in(0, 41.49].\\ 0.4 & \text{if } t\in(41.49, 64.53].\\ 0.6 & \text{if } t\in(64.53, 81.37].\\ 0.8 & \text{if } t\in(81.37, 93.26].\\ 1 & \text{if } t\in(93.26, 100].\end{cases}\quad 1-\nu_{A_4}(t)=\begin{cases}0 & \text{if } t=0.\\ 0.2 & \text{if } t\in(0, 37.43].\\ 0.4 & \text{if } t\in(37.43, 60.04].\\ 0.6 & \text{if } t\in(60.04, 77.53].\\ 0.8 & \text{if } t\in(77.53, 90.98].\\ 1 & \text{if } t\in(90.98, 100].\end{cases}$$

$$\mu_{A_5}(t)=\begin{cases}0 & \text{if } t=0.\\ 0.2 & \text{if } t\in(0, 49.24].\\ 0.4 & \text{if } t\in(49.24, 66.9].\\ 0.6 & \text{if } t\in(66.9, 82.61].\\ 0.8 & \text{if } t\in(82.61, 94.1].\\ 1 & \text{if } t\in(94.1, 100].\end{cases}\quad 1-\nu_{A_5}(t)=\begin{cases}0 & \text{if } t=0.\\ 0.2 & \text{if } t\in(0, 41.32].\\ 0.4 & \text{if } t\in(41.32, 64.89].\\ 0.6 & \text{if } t\in(64.89, 82.09].\\ 0.8 & \text{if } t\in(82.09, 93.49].\\ 1 & \text{if } t\in(93.49, 100].\end{cases}$$

Consider now the IF-sets A^+ and A^- defined in Equations(5.11) and (5.12), that are defined by $\mu_{A^+}=\mu_{A_2}$, $1-\nu_{A^+}=1-\nu_{A_1}$, $1-\nu_{A^-}=1-\nu_{A_5}$ and:

$$\mu_{A^-}=\begin{cases}0 & \text{if } t=0.\\ 0.2 & \text{if } t\in(0, 49.24].\\ 0.4 & \text{if } t\in(49.24, 69.81].\\ 0.6 & \text{if } t\in(69.81, 84.47].\\ 0.8 & \text{if } t\in(84.47, 94.56].\\ 1 & \text{if } t\in(94.56, 100].\end{cases}$$

Now, we consider two of the most usual measures of comparison of IF-divergences we can find in the literature, the Hausdorff and the Hamming distances that, as we have said in Section 5.1.3, are also local IF-divergences. Recall that they are defined, respectively, by:

$$d_H(A, B)=\int_0^{100}\max\{|\mu_A(\omega)-\mu_B(\omega)|, |\nu_A(\omega)-\nu_B(\omega)|\}\,d\omega.$$

$$l_{IFS}(A, B)=\frac{1}{2}\int_0^{100}|\mu_A(\omega)-\mu_B(\omega)|+|\nu_A(\omega)-\nu_B(\omega)|+|\pi_A(\omega)-\pi_B(\omega)|\,d\omega.$$

We represent the results on the next table.

	$l_{IFS}(A_i,A^+)$	$l_{IFS}(A_i,A^-)$	$d_H(A_i,A^+)$	$d_H(A_i,A^-)$
A_1	6.404	2.561	6.404	5.122
A_2	0.852	4.773	0.852	7.414
A_3	4.594	1.169	6.916	5.974
A_4	2.439	3.324	5.052	6.386
A_5	5.448	0.523	6.99	1.046

Now, we consider thre e different functions:

$$f_1(x, y) = y, \quad f_2(x, y) = -x \text{ and } f_3(x, y) = y - x.$$

f_1 only fo cus in the closest IF-set to the least optimal alternative; f_2 only fo cus in the closest IF-set tothe most optimal alte rn ative,while f_3 fo cus in the IF-set that is both closer to the most optimal alternative andless closer IF-set tothe leastoptimal alternative. We obtainthefollowing results:

l_{IFS}	f_1	f_2	f_3
A_1	2.561	−6.404	−3.843
A_2	4.773	−0.852	3.921
A_3	1.169	−4.594	−3.425
A_4	3.324	−2.439	0.885
A_5	0.523	−5.448	−4.925

d_{H}	f_1	f_2	f_3
A_1	5122	−6404	1282
A_2	7414	−0852	6562
A_3	5974	−6916	−0 942
A_4	6386	−5052	1334
A_5	1046	−6 99	−5 944

In the three cases, and with both IF-divergences, the preferred group is the second, that is, thegroup ofNordic countries. Theworstalternative, exceptfortheIF-divergence l_{IFS} and the function f_2, is the group A_5, that is the group of o ceanic countries. This means that the group of countries that has a b ette r wealth distribution is the group of Nordic countries, while thegreaterwealthinequalities are, in the most cases, in the group of oceanic countries.

5.4 Conclusions

The comparison of fuzzy sets is a topic that has been widely investigated, an several pap ers with mathematical theories can be found in the literature. However, when we move towards IF-sets th e efforts are somewhat scattered, and there is not an axiomatic approach to the comparisonof this kind ofsets.

For this reason we have develop ed a mathematical theory of the comparison of IF-sets. In particular, we have fo cused on IF-divergences,which are more restrictive measures than IF-dissi milarities. In particular, IF-divergences with the lo cal prop erty, named lo calIF-divergences, played an imp ortant role. As was exp ected,a c on nection between divergences for fuzzy sets and IF-divergences can be established, and we have found the conditions under which the lo cal prop erty, among other interesting prop erties, are preserved when we move from IF-divergences to divergences,and conversely, from divergences to IF-divergenc es.We also showed that these measures can be applied in pattern recognition and decision making, showing several examples.

On the other hand, we have investigated the connection between IVF-sets and Imprecise Probabilities. In this sense, we assumed thatthe IVF-set is defined ona probability space, and then it can be interpreted as a random interval. Then, we have investigated the probabilistic information enco ded by the random interval or its measurable selections, and we found conditions under which this probabilistic information coincides with the probabilistic information given by its asso ciated set of probabiliti es dominated by the upp er probability. We als o investigated the connection b etween our approach and other ones that can be found in the literature. In particular, the definition of probability for IF-sets given by Grzegorzewski and Mrowka is contained as a particular case of our theory.

The connection between IVF-sets and Imprecise Probabilities has allowed us to extend sto chastic dominance to the comparison of more than two p-b oxes simultaneously, determining also a complete relationship (i.e., avoiding incomparability). This metho d, that dep ends on the chosen IF-divergence, gives us a ranking of the p-b oxes. We have illustrated its behaviour continuing with the example of Section 4.4.1 in which we compare sets of Lorenz Curves.

For future research, some op en problems arise in the topic of comparison of IF-sets. On the one hand, it is p ossible to investigate under which conditions IF-divergences, and in particular lo cal IF-divergences, can define an entropy for IF-sets ([29]). On the other hand, as could be seen in the applications of IF-divergence, it is interesting to intro duce weights in the elements of the universe. In this situation it would be interesting to define lo cal IF-divergence with weights, and trying to find an analogous result to Theorem 5.29 to characterize them. Furthermore, we could investigate if it is possible to define lo cality with an op erator different than the sum; a t-conorm for instance. Moreover, our aim is to extend the lo cal prop erty to general universes, non-necessarily finite. With resp ect to the connection between IF-sets, IVF-sets and Imprecise Probabilities, we pretend to continue studying IF-sets and IVF-sets as bip olar mo dels for representing positive and negative information ([72, 73]).

Conclusiones y trabajo futuro

A lo largo de esta memoria se ha tratado el problema de la comparación de alternativas bajo ciertos tipos de falta de información: incertidumbre e imprecisión. La incertidumbre se refiere a situaciones en las que los posibles resultados del experimento están perfectamente descritos, pero el resultado del mismo es desconocido. Por otra parte, la imprecisión se refiere a situaciones en las que el resultado del experimento es conocido pero no es posible describirlo con precisión. Las herramientas utilizadas para modelar la incertidumbre y la imprecisión han sido la Teoría de las Probabilidades y la Teoría de los Conjuntos Intuicionísticos, respectivamente, mientras que la Teoría de las Probabilidades Imprecisas se ha utilizado para modelar ambas faltas de información simultáneas.

Cuando las alternativas a comparar están definidas bajo incertidumbre, éstas se han modelado mediante variables aleatorias, que son habitualmente comparadas mediante órdenes estocásticos. En esta memoria se han considerado, principalmente, dos de estos órdenes: la dominancia estocástica y la preferencia estadística. El primero de ellos es el orden estocástico más habitual en la literatura, y ha sido utilizado en diferentes ámbitos con destacables resultados. Por otra parte, la preferencia estadística es el método más adecuado para comparar variables cualitativas.

A pesar de que la dominancia estocástica es un método que ha sido investigado por varios autores, la preferencia estadística no ha sido estudiada con tanta profundidad. Ésta es la razón por la cual hemos estudiado sus propiedades como orden estocásticos. Uno de los resultados más destacados en este estudio es la relación de este método con la mediana. Esto demuestra que, mientras que la dominancia estocástica está relacionada con la media, la preferencia estadística es más cercana a otro parámetro de localización.

También hemos investigado la relación entre la dominancia estocástica y la preferencia estadística, y hemos encontrado condiciones bajo las cuales la dominancia estocástica de primer orden implica la preferencia estadística. Dado que la preferencia estadística depende de la distribución conjunta de las variables y, por tanto, de la cópula que las liga, dichas condiciones están también relacionadas con la cópula. El Teorema 3.64 resume estas condiciones: variables aleatorias independientes, variables aleatorias continuas ligadas por una cópula Arquimediana o variables aleatorias o bien continuas o bien discretas con soportes finitos que son comonótonas o contramonótonas. Además, hemos comprobado que esta relación no se cumple en general. Por tanto, de manera natural surge la siguiente cuestión: ¿es posible caracterizar las cópulas que hacen que la dominancia estocástica de primer orden implique la preferencia estadística?

Cuando las variables a comparar pertenecen a la misma familia paramétrica de distribuciones, como por ejemplo Bernoulli, exponencial, uniforme, Pareto, beta o normal, hemos visto que la dominancia estocástica y la preferencia estadística coinciden, y de hecho, ambos métodos se reducen a la comparación de sus esperanzas. Por esta razón es posible plantearse la siguiente conjetura: cuando las variables a comparar siguen la misma distribución perteneciente a la familia exponencial de distribuciones, tanto la dominancia estocástica como la preferencia estadística se reducen a la comparación de esperanzas y son, por tanto, equivalentes. Aunque éste es un problema abierto, una primera aproximación basada en simulaciones se ha realizado en [32].

La dominancia estocástica y la preferencia estadística son métodos de comparación de variables aleatorias por pares. Esto hace que en ocasiones no sean métodos adecuados para comparar más de dos variables simultáneamente. De hecho, la preferencia estadística es una relación no transitiva, y por lo tanto puede producir resultados ilógicos. Ésta es la razón que nos ha llevado a definir una generalización de la preferencia estadística para la comparación de más de dos variables simultáneamente. Siguiendo la misma aproximación que en el caso de la preferencia estadística, nuestra generalización da un grado de preferencia a cada una de las variables de manera que todos los grados sumen uno. Por lo tanto, la variable preferida será aquella con el mayor grado de preferencia. Para este método hemos estudiado su conexión con los órdenes estocásticos por pares. En particular, hemos visto que las mismas condiciones del Teorema 3.64 permiten asegurar que si una de las variables domina estocásticamente de primer grado al resto, entonces ésta es también preferida a todas las demás utilizando nuestra generalización de la preferencia estadística.

A la preferencia estadística general le podemos dar la siguiente interpretación: dado un conjunto de alternativas (en este caso variables aleatorias) tenemos que elegir entre la preferida, y podemos asignar a cada variable un grado de preferencia. Este grado de preferencia puede entenderse como cuánto de preferida es cada alternativa sobre el resto. Esto hace que la preferencia estadística general se pueda ver como una función de elección difusa ([81,207]). Un punto abierto sería por tanto estudiar la preferencia estadística general como una función de elección difusa.

Hay situaciones en las cuales las alternativas a comparar están definidas tanto bajo incertidumbre como bajo imprecisión. En tales casos, las variables aleatorias no recogen todas la información. En esta situación hemos modelado las alternativas mediante conjuntos de variables aleatorias con una interpretación epistémica: cada conjunto contiene la variable aleatoria original, que es desconocida. De cara a comparar estos conjuntos de alternativas, hemos tenido que extender los órdenes estocásticos para la comparación de conjuntos de variables aleatorias. Esta extensión da lugar a seis posibles métodos de ordenación de conjuntos de variables aleatorias. Una vez investigadas estas extensiones, nos hemos centrado en los casos en los que el orden estocástico utilizado es o bien la dominancia estocástica o bien la preferencia estadística, y hemos llamado a sus extensiones dominancia estocástica imprecisa y preferencia estadística imprecisa. La Proposición 4.19 y el Corolario 4.22 muestran que la dominancia estocástica imprecisa está relacionada con la comparación de las p-boxes asociadas a los conjuntos de variables aleatorias por medio de la dominancia estocástica. Estos resultados también nos permiten ver el estudio realizado por Denoeux ([61]) como un caso particular de nuestro estudio. Denoeux consideró dos medidas de creencia, y sus medidas de plausibilidad asociadas, y utilizó la dominancia estocástica para compararlas. Sin embargo, dado que las medidas de creencia y plausibilidad definen conjuntos de probabilidades, es posible compararlas mediante la dominancia estocástica imprecisa.

Lo mismo ocurre con posibilidades: una medida de posibilidad define un conjunto de probabilidades, y por lo tanto es posible utilizar la dominancia estocástica imprecisa para compararlas. En la Proposición 4.52 hemos dado una caracterización de la dominancia estocástica imprecisa para medidas de posibilidad con distribuciones de posibilidad continuas. Aquí surge un nuevo problema abierto: en caso de que las distribuciones de posibilidad asociadas a las distribuciones de posibilidad no sean continuas, ¿se cumple la misma caracterización de la Proposición 4.52?

Dos situaciones habituales dentro de la Teoría de la Decisión se pueden modelar mediante la comparación de conjuntos de variables aleatorias. Por una parte, hemos considerado la comparación de dos variables aleatorias con imprecisión en las utilidades. Esta falta de información ha sido modelada con conjuntos aleatorios. La información probabilística de un conjunto aleatorio se recoge en sus selecciones medibles. Por tanto, la comparación de conjuntos aleatorios se realiza mediante la comparación de sus conjuntos de selecciones medibles. Por otra parte, hemos considerado la comparación de variables aleatorias definidas sobre un espacio probabilístico donde la probabilidad no está definida de manera precisa. En esta situación, en vez de haber una única probabilidad, hemos considerado un conjunto de probabilidades. De esta manera también es posible definir dos conjuntos de variables aleatorias que recogen la información disponible. Para estas dos situaciones hemos investigado en particular las propiedades de la dominancia estocástica imprecisa y la preferencia estadística imprecisa, estudiando sus conexiones con la Teoría de las Probabilidades Imprecisas.

La preferencia estadística es un orden estocástico que está basado en la distribución conjunta de las variables aleatorias. El Teorema de Sklar asegura que la función de distribución conjunta de dos variables se puede expresar a través de las marginales mediante el uso de la cópula adecuada. Ahora bien, dados dos variable aleatorias definidas en un espacio de probabilidad descrito de manera imprecisa, el Teorema de Sklar no permite construir la distribución conjunta. Para tratar este problema, hemos investigado las p-boxes bivariantes y su conexión con las probabilidades inferiores coherentes. En particular, hemos visto que las funciones de distribución inferior y superior asociadas a un conjunto de funciones de distribución bivariantesno sonen general funciones de distribución bivariantes, puesto que no cumpl en la desigualdad de los rectángulos. Sin embargo, hemos visto que p ermiten defi nir una probabilidad inferior coherente, y a partir de resultados conocidos, las funciones de distribución inferior y sup erior cumplen cuatro desigualdades, llamadas (I-RI1), (I-RI2), (I-RI3) y (I-RI4), que puede n verse como las versiones imprecisas dela desigualdad de los rectángulos. La Prop osición 4.102 asegura que dos funciones de distribuci ón bivariantes, normalizadas y ordenadas define n una probabilidadinferiorcoherentecuandounadelasfuncionesde distribución estádefinida sobre un espacio binario. Como traba jo futuro, deseamos estudiarsi esta propiedadse cumple para funciones de distribución definidas sobre to do tip o de espacios, no necesariamente binarios.

El estudio de las p-b oxes bivariantes nos han permitido demostrar una versión imprecisa del Teorema de Sklar. Ennuestroestudiohemos asumidoque partimosde dos distribuciones margi nales imprecisas, definidas mediante p-b oxes, y de unconjunto de cópulas. En esta situación es p osible definir una p-b ox bivariante que defina a su vez una probabilidad inferior coherente. Además, hemosvisto queel recípro co no se cumple en general, puesto que una p-b ox bivariante que define una probabilidad inferior coherente no puede ser expresada, en general, a través de las p-b oxes marginales. Hemos comprobado que esta versión imprecisa del Te orema de Sklar es muy útil cuando hay que utilizar órdenes estocásticos ba jo imprecisión.

La extensión de los órdenes esto cásticos para la comparación de conjuntos de variables aleatorias tiene varias aplicaciones. Ademásdelasaplicacioneshabituales de los órdenes esto cásticos en la Teoría de la Decisión, hemos visto que también pueden ser aplicados a la comparación de Curvas de Lorenz aso ciadas a distintos grup os de países o regiones. Estos conjuntos de Curvas de Lorenz han sido comparados mediante la dominancia esto cástica imprecisa. Un estudio si milar se ha realizado para comparar tasas de sup ervivencia aso ciadas a distintos tip os cáncer, estudiando qué tip o de cáncer tiene peor diagnóstico.

Las alternativas definidas ba jo imprecisión, sin incertidumbre, se han mo delado mediante conjuntos intuicionísticos (IF-sets). IF-sets son un tip o de conjuntos que sirven para modelar información bip olar: considera los grados de p ertenencia y no pertenencia. Varios e je mplos de medidas de comparación de IF-sets se pueden encontrar en la literatura. Sin embargo, hasta este momento no se había desarrollado una teoría matemática. Por esta razón hemos considerados diferentes tip os de medidas de comparación, IF-disimilaridades, IF-divergencias, IF-disimilitudes y distancias, y las hemos estudiado de sde un punto de vis ta teórico. Por una parte hemos estudiado las relaciones existentes entre estas medidas, y hemos definido una medida general de comparación de IF-sets quecontiene alas otras medidas como casos particulares. Posteriorme nte nos hemos centrado en el estudiode las IF-divergencias, estudiando sus propiedades más interesantes. Enparticular, hemosconsideradounaclasedeIF-divergenciasquesatisface una condición de lo calidad. Tambiénhemos vistoquéconexión existeentre lasdivergencias para conjuntos difusos y las IF-divergencias. Por último, se han explicado posibles aplicaciones de las IF-divergencias en el reconocimiento de patrones y en la Teoría de la Decisión.

Pasamos a comentar algunos problemas abiertos relacionados con las IF-divergencias. Por una parte, en caso de que los elementos del espacio inicial tengan unos pesos asociados, parece posible extender las IF-divergencias locales considerando los pesos. Por otra parte, las IF-divergencias se podrían estudiar como entropías para IF-sets. Además, creemos que es posible extender la propiedad de localidad para universos no finitos, o incluso dar una defición de localidad basada en un operador diferente de la suma, como podría ser una t-conorma.

En las últimas fechas varios investigadores han centrado su atención en cómo las probabilidades imprecisas pueden modelar la información bipolar. Dado que los IF-sets también son utilizados en este mismo contexto, hemos establecido una conexión entre ambas teorías. Para ello, hemos considerado IF-sets definidos en un espacio probabilístico, y si entendemos los IF-sets como conjuntos intervalo-valorados, pueden ser vistos como conjuntos aleatorios. En esta situación, la información probabilística está recogida en el conjunto de selecciones medibles. Hemos visto condiciones bajo las cuales esta información coincide con la información probabilística dada por el conjunto credal asociado al conjunto aleatorio. Además, hemos visto que aproximaciones que ya se encontraban en la literatura se pueden ver como casos particulares de nuestro estudio.

La conexión entre los IF-sets y las probabilidades imprecisas nos han permitido extender la dominancia estocástica para la comparación de más de dos p-boxes al mismo tiempo. Como trabajo futuro, pensamos que este estudio podría ser completado. En particular, se podría estudiar la relación del procedimiento que hemos explicado con el uso de la habitual distancia de Kolmogorov entre funciones de distribución. Sin embargo, creemos que éste puede verse como un caso particular de nuestro estudio.

Conclusions and further research

This memory has dealt with the problem of comparing alternatives under lack of information. This lack of information can be of different kinds, and here we have assumed that it corresponds to either uncertainty or imprecision. Uncertainty refers to situations where the possible results of the experiment are precisely described, but the exact result is unknown; on the other hand, imprecision refers to situations in which the result of the experiment is known but it cannot be precisely described. In order to model uncertainty and imprecision we have used Probability Theory and Intuitionistic Fuzzy Set Theory, respectively; when both these features appear together in the decision problem, we have used the Theory of Imprecise Probabilities.

When the alternatives are subject to uncertainty in the outcomes, we we have modelled them as random variables, and have used stochastic orders so as to make a comparison between them. We have focused mainly in two different stochastic orders: stochastic dominance and statistical preference. The former is one of the most widely used stochastic orders we can find in the literature and the latter is of particular interest when comparing qualitative variables. Indeed, although stochastic dominance is a well-known method that has been widely investigated by several authors, statistical preference remained partly unexplored. For this reason we have studied several properties of this stochastic order. Possibly the most important one is its characterization in terms of the median, that serves us to compare it as a robust alternative to stochastic dominance, which is related to another location parameter: the mean.

We have also investigated the relationship between stochastic dominance and statistical preference, and we have found conditions under which (first degree) stochastic dominance implies statistical preference. Since statistical preference depends on the copula that links the variables into a joint distribution, the conditions we have obtained are also related to the copula. Theorem 3.64 summarizes such conditions: independent random variables, continuous random variables coupled by an Archimedean copula and either continuous or discrete random variables with finite supports that are either comonotonic or countermonotonic. In addition, we have also showed that the implication between these two stochastic orders does not hold in general. Thus, the first open question naturally arises: it is possible to characterize the set of copulas that makes first degree sto chastic dominance to imply statistical preference?

When the random variables to be compared belong to the same parametric family of distributions, like forinstance Bernoulli, exp onential, uniform, Pareto, beta or normal, we have seen that b oth sto chastic dominance and statistical preference coincide, and in fact, they are equivalent to compare the exp ectations of the random variables. This makes us to conjecture that when comparing two random variables that belong to the same parametric family of distribution within the exp onential family, then sto chastic dominance and statis tical preference reduce to the comparison of the exp ectations. Although this problem is still op en, a first ap proach, based onsimulations, has already b e done in [32].

Sto chastic dominance and statistical preference are pairwise metho ds of comparison of random variables. In this resp ect, they were not defined to compare more than two variables simultaneously. In fact, statistical preference is not atransitive relation, and therefore it may pro duce nonsensical results. For this reason we have gen eral ized statistical preference to the comparison of more than two random variables at th e same time. Withsimilarunderlyingideas tothoseof statistical preference, our generalization assigns apreference degree to any of the random variables, and the sum of these preference degrees is one.Then, the preferred randomvariableis theone with greater preference de gree.Forthisnew approachwehave investigated itsconnectionto the usual statistical preference and sto chastic dominance. In fact, the sameconditions of Theorem 3.64 that guarantee that sto chastic dominance implies statistical preference also ass ures that if there is a random variable that pairwise dominates all the others with resp ect to sto chastic dominance,then such random variable will be the preferred one with resp ect to our generalization of statistical preference.

A future line of research appears asso ciated with this general statistical preference. Given aset of alternatives (inthis case, random variables)out of which we have to cho ose the preferred one, wecanassign a degree of preference, that weunderstand as the strength of the preference of eachalternativeover the other. Then, the general statistical preference can b e seen as a fuzzy choice function de fined on a set of alternatives ([81, 207]). Thus, it may be interesting to investigate the prop erties of the general statistical preference inthe frameworkof fuzzy choice functions.

Onthe other hand, there are situation in which the alternatives to be compared are defined, notonlyunder uncertainty, butalso under imprecision. In such cases, randomvariables do notcollect allthe available information. Thus, we have mo delled the alternatives by me ans of sets of random variables with an epistemic interpretation:each set contains thereal unknownrandom variable. Inordertocomparethesesets, weneedto extend sto chastic orders to this general framework. In orde r to do this, we have considered any binary relation defined for the comparisonofsingle random variablesand wehave extended it for the comparison of sets of random variables. We have thus cons idered six p ossib le ways of ordering se ts of random variables. Afterinvestigating somegeneral prop erties of these extensions,we have fo cused in the cases in which binary relation is either stochastic dominance or statistical preference. We have called their extensions imprecise stochastic dominance and imprecise statistical preference. Proposition 4.19 and Corollary 4.22 showed that the former is clearly connected to the comparison of the bounds of the associated p-boxes by means of stochastic dominance. These results also helped to show that the approach given by Denoeux ([61]) is a particular case of our more general framework. Denoeux considered two belief functions, and their respective plausibility functions, and used stochastic dominance to compare them. Since each belief and plausibility function can be represented as a set of probabilities, and therefore imprecise stochastic dominance can be applied; we have seen that our definitions become the ones given by Denoeux for this particular case.

The same happens with possibilities: each possibility defines a set of probabilities, and therefore the imprecise stochastic dominance can be used to compare them. Proposition 4.52 showed a characterization of the imprecise stochastic dominance for possibility measures with continuous possibility distribution. Thus, an open problem is to investigate if such characterization also holds for possibility measures with non-continuous possibility distributions.

We have explored two situations that are usually present in decision making and that can be modelled by means of the comparison of sets of random variables. On the one hand, we have considered the comparison of two random variables with imprecision on the utilities. We have modelled this imprecision with random sets. Since under our epistemic interpretation the set of measurable selections of a random set encodes its probabilistic information, the comparison of random sets must be made by means of the comparison of their associated credal sets. On the other hand, we can also compare random variables defined on a probability space with a non-precisely determined probability; in that case, we have to consider a set of probabilities instead a single one. In this situation we can also consider two sets of random variables that summarise all the available information. For these two particular situations we have explored the properties of imprecise stochastic dominance and statistical preference, and we heave investigated their connection to imprecise probabilities.

We know that statistical preference is a stochastic order that is based on the joint distribution of the random variables. By Sklar's Theorem, this joint distribution is determined combining the marginals by means of a copula. However, given two random variables defined in a probability space with imprecise beliefs, Sklar's Theorem does not allow to define the joint distribution. In order to solve this problem, we have investigated bivariate p-boxes and how they can define a coherent lower probability. In particular, we have seen that the lower and upper distributions associated with a set of bivariate distribution functions are not in general bivariate distribution functions because they violate the rectangle inequality. However, we have seen that they define a coherent lower probability and they satisfy four inequalities, named (I-RI1), (I-RI2), (I-RI3) and (I-RI4), that can be seen as the imprecise versions of the rectangle inequality. We have seen in Proposition 4.102 that given two ordered normalized bivariate distribution functions that satisfy them, they define a coherent lower probability if one of the normalized functions is defined on a binary space. An open problem for future research is to investigate if this property also holds for normalized functions defined on any space.

The study of bivariate p-boxes have allowed to define an imprecise version of Sklar's Theorem. We have assumed that we have two imprecise marginal distributions, that we model by means of p-boxes, and we have a set of possible copulas that link them. In this situation it is possible to define a bivariate p-box that defines a coherent lower probability. However, the second part of the Sklar's Theorem does not hold, because a bivariate p-box that defines a coherent lower probability cannot be expressed, in general, by means of the marginal p-boxes. We have also seen how this imprecise version is very useful when dealing with bivariate stochastic orders with imprecision.

The extension of stochastic orders to the comparison of sets of random variables we have proposed has several applications. Besides the usual application of stochastic orders in decision making, we have seen that they can be also applied to the comparison of the inequality indices between groups of countries. In this work, we have considered the Lorenz curve of each country, that measures the inequality of such country, and we have grouped them by geographical areas. Then, we have compared these groups of Lorenz curves using the imprecise stochastic dominance. We have made a similar approach to the comparison of cancer survival rates, grouping them by cancer sites, and we have analyzed which cancer site has a worst prognosis.

Alternatives defined under imprecision, without uncertainty, have been modelled by means of IF-sets. IF-sets are bipolar models that allow to define membership and non-membership degrees. Several examples of measures of comparison of IF-sets had been proposed in the literature. However, a mathematical theory had not been developed. For this reason we have considered different kinds of measures, IF-dissimilarities, IF-divergences, IF-dissimilitudes and distances, and we have investigated them from a theoretical point of view. First of all, we have seen the relationships between these measures, and we have defined a general measure of comparison of IF-sets that contains them as particular cases. Then, we have focused on IF-divergences and we have investigated its main properties. In particular, we have considered one instance of IF-divergences, those that satisfy a local property. We have also seen the connection between IF-divergences and divergences for fuzzy sets. We have also showed how IF-divergences can be applied within pattern recognition and decision making.

There are several open problems related to this study of IF-divergences. On the one hand, it would be interesting to define local IF-divergences that take into account a weight function on the the elements of the initial space. On the other hand, IF-divergences could be studied as entropies for IF-sets. Furthermore, it is possible to extend the local property to spaces non-necessarily finite, and also to define the local property by means of an operator different than the sum, like t-conorms, for instance.

Currently, several authors have been investigating how imprecise probabilities can be used to model bipolar information. Since IF-sets are also useful in this context, we have established a connection between both theories. We have assumed that IF-sets are defined in a probability space; if we understand them as IVF-sets, they can then be seen as random sets. In that case, their probabilistic information can be encoded by the set of measurable selections. We have seen conditions under which such information coincides with the probabilistic information given the credal set associated to the random set. Furthermore, we have seen how previous approaches made for defining a probability measure on IF-sets can be embedded into our approach.

The connection between IF-sets and imprecise probabilities has allowed us to extend stochastic dominance to the comparison of more than two p-boxes simultaneously. For future research, we think that this proposal could be studied more thoroughly. For instance, a similar extension of stochastic dominance may be made by using the usual Kolmogorov distance between cumulative distribution functions. It would be interesting to determine if this becomes a particular case of our more general framework.

A Appendix: Basic Results

In this Appendix we prove some results that we have used throughout this report.

Lemma A.1 *Let a, b and c be three real numbers in $[0,1]$. Then*

a) $\max\{0, \min\{a,c\} - \min\{b,c\}\} \leq \max\{0, a-b\}$ *and*
$\max\{0, \max\{a,c\} - \max\{b,c\}\} \leq \max\{0, a-b\}$.
b) $\max(|\max\{a,c\} - \max\{b,c\}|, |\min\{a,c\} - \min\{b,c\}|) \leq |a-b|$.

Proof We distinguish the following cases, depending on the minimum and the maximum of $\{a,c\}$ and $\{b,c\}$:

1. Assume that $\min\{a,c\} = a$ and $\min\{b,c\} = b$, and consequently $\max\{a,c\} = \max\{b,c\} = c$. Then:

   a) $\max\{0, \min\{a,c\} - \min\{b,c\}\} = \max\{0, a-b\}$.
   $\max\{0, \max\{a,c\} - \max\{b,c\}\} = 0 \leq \max\{0, a-b\}$.
   b) $|\max\{a,c\} - \max\{b,c\}| = |c-c| = 0 \leq |a-b|$.
   $|\min\{a,c\} - \min\{b,c\}| = |a-b|$.

2. Assume next that $\min\{a,c\} = a$ and $\min\{b,c\} = c$, and therefore $\max\{a,c\} = c$ and $\max\{b,c\} = b$. Note that, since $\min\{a,c\} = a$, then $a \leq c$, and therefore $a - c \leq 0$. Moreover, it also holds that $c \leq b$, and consequently $a \leq c \leq b$. Hence:

   a) $\max\{0, \min\{a,c\} - \min\{b,c\}\} = \max\{0, a-c\} = 0$
   $\leq \max\{0, a-b\}$.
   $\max\{0, \max\{a,c\} - \max\{b,c\}\} = \max\{0, c-b\} = 0$
   $\leq \max\{0, c-b\}$.
   b) $|\max\{a,c\} - \max\{b,c\}| = |c-b| \leq |a-b|$.
   $|\min\{a,c\} - \min\{b,c\}| = |a-c| \leq |a-b|$.

3. Thirdly, assume that $\min\{a,c\} = c$ and $\min\{b,c\} = b$, whence $\max\{a,c\} = a$ and $\max\{b,c\} = c$. In such a case, $c \le a$ and $b \le c$, and therefore $b \le c \le a$, that implies $c - b \le a - b$ and $a - c \le a - b$. Hence:

$$\begin{aligned}
a)\quad & \max\{0, \min\{a,c\} - \min\{b,c\}\} = \max\{0, c - b\} \\
& \qquad \le \max\{0, a - b\}. \\
& \max\{0, \max\{a,c\} - \max\{b,c\}\} = \max\{0, a - c\} \\
& \qquad \le \max\{0, a - b\}. \\
b)\quad & |\max\{a,c\} - \max\{b,c\}| = |a - c| \le |a - b|. \\
& |\min\{a,c\} - \min\{b,c\}| = |c - b| \le |a - b|.
\end{aligned}$$

4. Finally, assume that $\min\{a,c\} = \min\{b,c\} = c$, and consequently $\max\{a,c\} = a$ and $\max\{b,c\} = b$. Then:

$$\begin{aligned}
a)\quad & \max\{0, \min\{a,c\} - \min\{b,c\}\} = 0 \le \max\{0, a - b\}. \\
& \max\{0, \max\{a,c\} - \max\{b,c\}\} = \max\{0, a - b\}. \\
b)\quad & |\max\{a,c\} - \max\{b,c\}| = |a - b|. \\
& |\min\{a,c\} - \min\{b,c\}| = |c - c| = 0 \le |a - b|. \qquad \blacksquare
\end{aligned}$$

Lemma A.2 *If (a_1,a_2), (b_1,b_2) and (c_1,c_2) are elements on $T = \{(x,y) \in [0,1]^2 \mid x + y \le 1\}$, it holds that:*

$$\begin{aligned}
\alpha = {} & |a_1 - b_1| + |a_2 - b_2| + |a_1 + a_2 - b_1 - b_2| \\
\ge {} & |\max\{a_1,c_1\} - \max\{b_1,c_1\}| + |\min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
& + |\max\{a_1,c_1\} + \min\{a_2,c_2\} - \max\{b_1,c_1\} - \min\{b_2,c_2\}| = \beta.
\end{aligned}$$

Proof Let us consider the following possibilities:

1. $a_1,b_1 \le c_1$ and $a_2,b_2 \le c_2$. Then:

$$\begin{aligned}
\beta = {} & |c_1 - c_1| + |a_2 - b_2| + |c_1 + a_2 - c_1 - b_2| = 2\,|a_2 - b_2| \\
\le {} & |a_2 - b_2| + |a_1 - b_1| + |a_1 + a_2 - b_1 - b_2| = \alpha.
\end{aligned}$$

2. $a_1,b_1 \le c_1$ and $c_2 \le a_2,b_2$. Then it holds that:

$$\beta = |c_1 - c_1| + |c_2 - c_2| + |c_1 + c_2 - c_1 - c_2| = 0 \le \alpha.$$

3. $a_1,b_1 \le c_1$ and $b_2 \le c_2 \le a_2$:

$$\begin{aligned}
\beta = {} & |c_1 - c_1| + |c_2 - b_2| + |c_1 + c_2 - c_1 - b_2| = 2\,|c_2 - b_2| \\
\le {} & 2|a_2 - b_2| \le \alpha.
\end{aligned}$$

4. $c_1 \le a_1,b_1$ and $c_2 \le a_2,b_2$:

$$\begin{aligned}
\beta = {} & |a_1 - b_1| + |c_2 - c_2| + |a_1 + c_2 - b_1 - c_2| = 2\,|a_1 - b_1| \\
\le {} & |a_1 - b_1| + |a_2 - b_2| + |a_1 + a_2 - b_1 - b_2| = \alpha.
\end{aligned}$$

5. $c_1 \le a_1, b_1$ and $b_2 \le c_2 \le a_2$.

$$
\begin{aligned}
\beta &= |a_1 - b_1| + |c_2 - b_2| + |a_1 + c_2 - b_1 - b_2| \\
&= \begin{cases} |b_1 - a_1| + (c_2 - b_2) + (a_1 - b_1) - (b_2 - c_2) & \text{if } a_1 - b_1 \ge b_2 - c_2 \\ |b_1 - a_1| + (c_2 - b_2) + (b_2 - c_2) - (a_1 - b_1) & \text{if } a_1 - b_1 < b_2 - c_2 \end{cases} \\
&= \begin{cases} |b_1 - a_1| + (a_1 - b_1) + 2(c_2 - b_2) & \text{if } a_1 - b_1 \ge b_2 - c_2 \\ 2|b_1 - a_1| & \text{if } a_1 - b_1 < b_2 - c_2 \end{cases} \\
&\le \begin{cases} |b_1 - a_1| + (a_1 - b_1) + (c_2 - b_2) + (a_2 - b_2) & \text{if } a_1 - b_1 \ge b_2 - c_2 \\ |a_1 - b_1| + |a_2 - b_2| + |a_1 + a_2 - b_1 - b_2| & \text{if } a_1 - b_1 < b_2 - c_2 \end{cases} \\
&\le \begin{cases} |b_1 - a_1| + (a_1 - b_1) + (a_2 - b_2) + (a_2 - b_2) & \text{if } a_1 - b_1 \ge b_2 - c_2 \\ |a_1 - b_1| + |a_2 - b_2| + |a_1 + a_2 - b_1 - b_2| & \text{if } a_1 - b_1 < b_2 - c_2 \end{cases} \\
&\le \alpha.
\end{aligned}
$$

6. $b_1 \le c_1 \le a_1$ and $b_2 \le c_2 \le a_2$.

$$
\begin{aligned}
\beta &= |a_1 - c_1| + |c_2 - b_2| + |a_1 + c_2 - c_1 - b_2| \\
&= (a_1 - c_1) + (c_2 - b_2) + (a_1 - c_1) + (c_2 - b_2) \\
&= 2(a_1 - c_1) + 2(c_2 - b_2) \le 2(a_1 - b_1) + 2(a_2 - b_2) \le \alpha.
\end{aligned}
$$

7. $b_1 \le c_1 \le a_1$ and $a_2 \le c_2 \le b_2$.

$$
\begin{aligned}
\beta &= |a_1 - c_1| + |a_2 - c_2| + |a_1 + a_2 - c_1 - c_2| \\
&= \begin{cases} (a_1 - c_1) + (c_2 - a_2) + (a_1 - c_1) + (a_2 - c_2) & \text{if } a_1 - c_1 \ge c_2 - a_2 \\ (a_1 - c_1) + (c_2 - a_2) - (a_1 - c_1) - (a_2 - c_2) & \text{if } a_1 - c_1 < c_2 - a_2 \end{cases} \\
&= \begin{cases} 2(a_1 - c_1) \le 2(a_1 - b_1) & \text{if } a_1 - c_1 \ge c_2 - a_2 \\ 2(c_2 - a_2) \le 2(b_2 - a_2) & \text{if } a_1 - c_1 < c_2 - a_2 \end{cases} \\
&\le \begin{cases} 2(a_1 - b_1) & \text{if } a_1 - c_1 \ge c_2 - a_2 \\ 2(b_2 - a_2) & \text{if } a_1 - c_1 < c_2 - a_2 \end{cases} \le \alpha.
\end{aligned}
$$

In the remaining cases, it is enough to exchange the roles of (a_1, b_1), (a_2, b_2) and to apply the previous cases. ∎

Lemma A.3 *If (a_1, a_2), (b_1, b_2) and (c_1, c_2) are elements on $T = \{(x, y)\ [0, 1]^2 \mid x + y \le 1\}$, then it holds that:*

$$
\begin{aligned}
&|a_1 - b_1 - a_2 + b_2| + |a_1 - b_1| + |a_2 - b_2| \ge \\
&\quad |\max\{a_1, c_1\} - \max\{b_1, c_1\} - \min\{a_2, c_2\} + \min\{b_2, c_2\}| + \\
&\quad |\max\{a_1, c_1\} - \max\{b_1, c_1\}| + |\min\{a_2, c_2\} - \min\{b_2, c_2\}|.
\end{aligned}
$$

Proof Let us consider some cases.

1. $a_1, b_1 \le c_1$ and $a_2, b_2 \le c_2$.

$$\begin{aligned}
&|\max\{a_1,c_1\} - \max\{b_1,c_1\} - \min\{a_2,c_2\} + \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\}| + |\min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&\quad = |c_1 - c_1 - a_2 + b_2| + |c_1 - c_1| + |a_2 - b_2| = 2\,|b_2 - a_2| \\
&\quad \le |a_1 - b_1 - a_2 + b_2| + |a_1 - b_1| + |a_2 - b_2|.
\end{aligned}$$

2. $a_1, b_1 \le c_1$ and $c_2 \le a_2, b_2$.

$$\begin{aligned}
&|\max\{a_1,c_1\} - \max\{b_1,c_1\} - \min\{a_2,c_2\} + \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\}| + |\min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&\quad = |c_1 - c_1 - c_2 + c_2| + |c_1 - c_1| + |c_2 - c_2| = 0 \\
&\quad \le |a_1 - b_1 - a_2 + b_2| + |a_1 - b_1| + |a_2 - b_2|.
\end{aligned}$$

3. $a_1, b_1 \le c_1$ and $b_2 \le c_2 \le a_2$.

$$\begin{aligned}
&|\max\{a_1,c_1\} - \max\{b_1,c_1\} - \min\{a_2,c_2\} + \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\}| + |\min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&\quad = |c_1 - c_1 - c_2 + b_2| + |c_1 - c_1| + |c_2 - b_2| = 2\,|c_2 - b_2| \\
&\quad \le |a_1 - b_1 - a_2 + b_2| + |a_1 - b_1| + |a_2 - b_2|.
\end{aligned}$$

4. $a_1, b_1 \le c_1$ and $a_2 \le c_2 \le b_2$. It suffices to exchange the roles of (a_1, a_2) and (b_1, b_2) and to apply the previous case.

5. $c_1 \le a_1, b_1$ and $a_2, b_2 \le c_2$. Take (a_2, a_1) and (b_2, b_1) and apply case 2.

6. $c_1 \le a_1, b_1$ and $c_2 \le a_2, b_2$.

$$\begin{aligned}
&|\max\{a_1,c_1\} - \max\{b_1,c_1\} - \min\{a_2,c_2\} + \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\}| + |\min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&\quad = |a_1 - b_1 - c_2 + c_2| + |a_1 - b_1| + |c_2 - c_2| = 2\,|a_1 - b_1| \\
&\quad = |a_1 - b_1 - a_2 + b_2| + |a_1 - b_1| + |a_2 - b_2|.
\end{aligned}$$

7. $c_1 \le a_1, b_1$ and $b_2 \le c_2 \le a_2$.

$$\begin{aligned}
&|\max\{a_1,c_1\} - \max\{b_1,c_1\} - \min\{a_2,c_2\} + \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\}| + |\min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&= |a_1 - b_1 - c_2 + b_2| + |a_1 - b_1| + |c_2 - b_2| \\
&= \begin{cases} (a_1 - b_1) - (c_2 - b_2) + |a_1 - b_1| + (c_2 - b_2) & \text{if } a_1 - b_1 \ge c_2 - b_2 \\ 2(c_2 - b_2) - (a_1 - b_1) + |a_1 - b_1| & \text{if } a_1 - b_1 \le c_2 - b_2 \end{cases} \\
&\le \begin{cases} 2|a_1 - b_1| & \text{if } a_1 - b_1 \ge c_2 - b_2 \\ (a_2 - b_2) - (a_1 - b_1) + |a_1 - b_1| + |a_2 - b_2| & \text{if } a_1 - b_1 \le c_2 - b_2 \end{cases} \\
&= |a_1 - b_1| + |a_2 - b_2| + |a_1 - b_1 - a_2 + b_2|.
\end{aligned}$$

8. $c_1 \leq a_1, b_1$ and $a_2 \leq c_2 \leq b_2$. It suffices to exchange (a_1, a_2) and (b_1, b_2) and to apply the previouscase.

9. $b_1 \leq c_1 \leq a_1$ and $a_2, b_2 \leq c_2$. It is enough to consid er (a_2, a_1) and (b_1, b_2) and to apply case 3.

10. $b_1 \leq c_1 \leq a_1$ and $c_2 \leq a_2, b_2$. Itsuffices toconsider (a_2, a_1) and (b_1, b_2) and to apply case 7.

11. $b_1 \leq c_1 \leq a_1$ and $b_2 \leq c_2 \leq a_2$.

$$
\begin{aligned}
&|\max\{a_1, c_1\} - \max\{b_1, c_1\} - \min\{a_2, c_2\} + \min\{b_2, c_2\}| \\
&+ |\max\{a_1, c_1\} - \max\{b_1, c_1\}| + |\min\{a_2, c_2\} - \min\{b_2, c_2\}| \\
&= |a_1 - c_1 - c_2 + b_2| + |a_1 - c_1| + |c_2 - b_2| \\
&= \begin{cases} 2(a_1 - c_1) + (c_2 - b_2) - (c_2 - b_2) & \text{if } a_1 - c_1 \geq c_2 - b_2 \\ (a_1 - c_1) + 2(c_2 - b_2) - (a_1 - c_1) & \text{if } a_1 - c_1 \leq c_2 - b_2 \end{cases} \\
&\leq \begin{cases} 2(a_1 - b_1) & \text{if } a_1 - c_1 \geq c_2 - b_2 \\ 2(a_2 - b_2) & \text{if } a_1 - c_1 \leq c_2 - b_2 \end{cases} \\
&= |a_1 - b_1| + |a_2 - b_2| + |a_1 - b_1 - a_2 + b_2|.
\end{aligned}
$$

12. $b_1 \leq c_1 \leq a_1$ and $a_2 \leq c_2 \leq b_2$.

$$
\begin{aligned}
&|\max\{a_1, c_1\} - \max\{b_1, c_1\} - \min\{a_2, c_2\} + \min\{b_2, c_2\}| \\
&+ |\max\{a_1, c_1\} - \max\{b_1, c_1\}| + |\min\{a_2, c_2\} - \min\{b_2, c_2\}| \\
&\quad = |a_1 - c_1 - a_2 + c_2| + |a_1 - c_1| + |a_2 - c_2| \\
&\quad = 2(a_1 - c_1) + 2(c_2 - a_2) \\
&\quad \leq 2(a_1 - b_1) + 2(b_2 - a_2) \\
&\quad = |a_1 - b_1| + |a_2 - b_2| + |a_1 - b_1 - a_2 + b_2|.
\end{aligned}
$$

13. $a_1 \leq c_1 \leq b_1$. Itisenough toconsider (a_2, a_1) and (b_2, b_1) andto applythe previous cases. ∎

Lemma A.4 *If* (a_1, a_2), (b_1, b_2) *and* (c_1, c_2) *are elementson* $T = \{(x, y) \in [0, 1]^2 \mid x + y \leq 1\}$*, then:*

$$
\begin{aligned}
&|a_1 - b_1| + |a_2 - b_2| + |a_1 - b_1 - a_2 + b_2| + |a_1 + a_2 - b_1 - b_2| \geq \\
&\quad |\max\{a_1, c_1\} - \max\{b_1, c_1\}| + |\min\{a_2, c_2\} - \min\{b_2, c_2\}| + \\
&\quad |\max\{a_1, c_1\} - \max\{b_1, c_1\} - \min\{a_2, c_2\} + \min\{b_2, c_2\}| + \\
&\quad |\max\{a_1, c_1\} - \max\{b_1, c_1\} + \min\{a_2, c_2\} - \min\{b_2, c_2\}|.
\end{aligned}
$$

Pro of Throughout this proof we will use the fact that $|x + y| + |x - y| = \max\{2|x|, 2|y|\}$. Let us consider the following possibilities.

1. $a_1, b_1 \le c_1$ and $a_2, b_2 \le c_2$.

$$\begin{aligned}
&|\max\{a_1,c_1\} - \max\{b_1,c_1\}| + |\min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\} - \min\{a_2,c_2\} + \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\} + \min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&\quad = |c_1 - c_1| + |a_2 - b_2| + |c_1 - c_1 - a_2 + b_2| + |c_1 - c_1 + a_2 - b_2| \\
&\quad = 3|a_2 - b_2| \le |a_2 - b_2| + |a_2 - b_2 - a_1 + b_1| + |a_2 - b_2 + a_1 - b_1| \\
&\quad \le |a_1 - b_1| + |a_2 - b_2| + |a_2 - b_2 - a_1 + b_1| + |a_2 - b_2 + a_1 - b_1|.
\end{aligned}$$

2. $a_1, b_1 \le c_1$ and $c_2 \le a_2, b_2$.

$$\begin{aligned}
&|\max\{a_1,c_1\} - \max\{b_1,c_1\}| + |\min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\} - \min\{a_2,c_2\} + \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\} + \min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&= |c_1 - c_1| + |c_2 - c_2| + |c_1 - c_1 - c_2 + c_2| + |c_1 - c_1 + c_2 - c_2| \\
&= 0 \le |a_1 - b_1| + |a_2 - b_2| + |a_2 - b_2 - a_1 + b_1| + |a_2 - b_2 + a_1 - b_1|.
\end{aligned}$$

3. $a_1, b_1 \le c_1$ and $b_2 \le c_2 \le a_2$.

$$\begin{aligned}
&|\max\{a_1,c_1\} - \max\{b_1,c_1\}| + |\min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\} - \min\{a_2,c_2\} + \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\} + \min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&\quad = |c_1 - c_1| + |c_2 - b_2| + |c_1 - c_1 - c_2 + b_2| + |c_1 - c_1 + c_2 - b_2| \\
&\quad = 3|c_2 - b_2| \le 3|a_2 - b_2| \\
&\quad = |a_2 - b_2| + |a_2 - b_2 - a_1 + b_1| + |a_2 - b_2 + a_1 - b_1| \\
&\quad \le |a_1 - b_1| + |a_2 - b_2| + |a_2 - b_2 - a_1 + b_1| + |a_2 - b_2 + a_1 - b_1|.
\end{aligned}$$

4. $a_1, b_1 \le c_1$ and $a_2 \le c_2 \le a_2$. It suffi ces to exchange the roles of (a_1, a_2) and (b_1, b_2).

5. $c_1 \le a_1, b_1$ and $a_2, b_2 \le c_2$. It suffices toconsider (a_2, a_1) and (b_2, b_1) and to apply case 2.

6. $c_1 \le a_1, b_1$ and $c_2 \le a_2, b_2$.

$$\begin{aligned}
&|\max\{a_1,c_1\} - \max\{b_1,c_1\}| + |\min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\} - \min\{a_2,c_2\} + \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\} + \min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&\quad = |a_1 - b_1| + |c_2 - c_2| + |a_1 - b_1 + c_2 - c_2| + |a_1 - b_1 - c_2 + c_2| \\
&\quad = 3|a_1 - b_1| \le |a_1 - b_1| + |a_2 - b_2 - a_1 + b_1| + |a_2 - b_2 + a_1 - b_1| \\
&\quad \le |a_1 - b_1| + |a_2 - b_2| + |a_2 - b_2 - a_1 + b_1| + |a_2 - b_2 + a_1 - b_1|.
\end{aligned}$$

7. $c_1 \leq a_1, b_1$ and $b_2 \leq c_2 \leq a_2$.

$$\begin{aligned}
&|\max\{a_1,c_1\} - \max\{b_1,c_1\}| + |\min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\} - \min\{a_2,c_2\} + \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\} + \min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&\quad = |a_1 - b_1| + |c_2 - b_2| + |a_1 - b_1 - c_2 + b_2| + |a_1 - b_1 + c_2 - b_2| \\
&\quad = |a_1 - b_1| + |c_2 - b_2| + 2\max(|a_1 - b_1|, |c_2 - b_2|) \\
&\quad \leq |a_1 - b_1| + |a_2 - b_2| + 2\max(|a_1 - b_1|, |a_2 - b_2|) \\
&\quad \leq |a_1 - b_1| + |a_2 - b_2| + |a_2 - b_2 - a_1 + b_1| + |a_2 - b_2 + a_1 - b_1|.
\end{aligned}$$

8. $c_1 \leq a_1, b_1$ and $a_2 \leq c_2 \leq a_2$. It sufficesto exchange the roles of (a_1, a_2) and (b_1, b_2) and to applythe previouscase.

9. $b_1 \leq c_1 \leq a_1$ and $a_2, b_2 \leq c_2$. It is enough to consid er (a_2, a_1) and (b_2, b_1) and to apply case 3.

10. $b_1 \leq c_1 \leq a_1$ and $c_2 \leq a_2, b_2$. Consid er (a_2, a_1) and (b_2, b_1) and to app ly case 7.

11. $b_1 \leq c_1 \leq a_1$ and $b_2 \leq c_2 \leq a_2$.

$$\begin{aligned}
&|\max\{a_1,c_1\} - \max\{b_1,c_1\}| + |\min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\} - \min\{a_2,c_2\} + \min\{b_2,c_2\}| \\
&+ |\max\{a_1,c_1\} - \max\{b_1,c_1\} + \min\{a_2,c_2\} - \min\{b_2,c_2\}| \\
&\quad = |a_1 - c_1| + |c_2 - b_2| + |a_1 - c_1 - c_2 + b_2| + |a_1 - c_1 + c_2 - b_2| \\
&\quad = |a_1 - c_1| + |c_2 - b_2| + 2\max(|a_1 - c_1|, |c_2 - b_2|) \\
&\quad \leq |a_1 - b_1| + |a_2 - b_2| + 2\max(|a_1 - b_1|, |a_2 - b_2|) \\
&\quad = |a_1 - b_1| + |a_2 - b_2| + |a_2 - b_2 - a_1 + b_1| + |a_2 - b_2 + a_1 - b_1|.
\end{aligned}$$

12. $b_1 \leq c_1 \leq a_1$ and $a_2 \leq c_2 \leq b_2$. It suffices to exchange the roles of (a_1, a_2) and (b_1, b_2) and to appl y the previous case.

13. $a_1 \leq c_1 \leq b_1$. Itsuffices to exchange (a_1, a_2) and (b_1, b_2) and to apply the previous cases. ∎

Lemma A.5 *If* (a_1, a_2), (b_1, b_2) *and* (c_1, c_2) *are three elements in* $T = \{(x, y) \in [0, 1]^2 \mid x + y \leq 1\}$*, then:*

$$\begin{aligned}
&|\max\{a_1 - 0.5, 0\} - \max\{b_1 - 0.5, 0\}| + \\
&|\max\{a_2 - 0.5, 0\} - \max\{b_2 - 0.5, 0\}| \geq \\
&\quad |\max\{\max\{a_1, c_1\} - 0.5, 0\} - \max\{\max\{b_1, c_1\} - 0.5, 0\}| + \\
&\quad |\max\{\min\{a_2, c_2\} - 0.5, 0\} - \max\{\min\{b_2, c_2\} - 0.5, 0\}|.
\end{aligned}$$

Pro of In order to prove this result, we are going to prove the following inequalities:

$$\begin{aligned}
&|\max\{a - 0.5, 0\} - \max\{b - 0.5, 0\}| \geq \\
&\quad |\max\{\max\{a,c\} - 0.5, 0\} - \max\{\max\{b,c\} - 0.5, 0\}|, \\
&|\max\{a - 0.5, 0\} - \max\{b - 0.5, 0\}| \geq \\
&\quad |\max\{\min\{a,c\} - 0.5, 0\} - \max\{\min\{b,c\} - 0.5, 0\}|,
\end{aligned}$$

for every a, b, c $[0, 1]$. Let us consider several cases.

1. $a \le b \le c$.

$$
\begin{aligned}
&|\max\{\max\{a,c\} - 0.5, 0\} - \max\{\max\{b,c\} - 0.5, 0\}| \\
&\quad = |\max\{c - 0.5, 0\} - \max\{c - 0.5, 0\}| \\
&\quad = 0 \le |\max\{a - 0.5, 0\} - \max\{b - 0.5, 0\}|. \\
&|\max(\min\{a,c\} - 0.5, 0\} - \max\{\min\{b,c\} - 0.5, 0\}| \\
&\quad = |\max\{a - 0.5, 0\} + \max\{b - 0.5, 0\}|.
\end{aligned}
$$

2. $a \le c \le b$. This implies that $b - 0.5 \ge c - 0.5 \ge a - 0.5$, and therefore $\max\{b - 0.5, 0\} \ge \max\{c - 0.5, 0\} \ge \max\{a - 0.5, 0\}$.

$$
\begin{aligned}
&|\max\{\max\{a,c\} - 0.5, 0\} - \max\{\max\{b,c\} - 0.5, 0\}| \\
&\quad = |\max\{c - 0.5, 0\} - \max\{b - 0.5, 0\}| \\
&\quad \le |\max\{a - 0.5, 0\} - \max\{b - 0.5, 0\}| \\
&|\max\{\min\{a,c\} - 0.5, 0\} - \max\{\min\{b,c\} - 0.5, 0\}| \\
&\quad = |\max\{a - 0.5, 0\} - \max\{c - 0.5, 0\}| \\
&\quad \le |\max\{a - 0.5, 0\} - \max\{b - 0.5, 0\}|.
\end{aligned}
$$

3. $b \le a \le c$.

$$
\begin{aligned}
&|\max\{\max\{a,c\} - 0.5, 0\} - \max\{\max\{b,c\} - 0.5, 0\}| \\
&\quad = |\max\{c - 0.5, 0\} - \max\{c - 0.5, 0\}| \\
&\quad = 0 \le |\max\{a - 0.5, 0\} - \max\{b - 0.5, 0\}|. \\
&|\max\{\min\{a,c\} - 0.5, 0\} - \max\{\min\{b,c\} - 0.5, 0\}| \\
&\quad = |\max\{a - 0.5, 0\} - \max\{b - 0.5, 0\}|.
\end{aligned}
$$

4. $b \le c \le a$. Then $a - 0.5 \ge c - 0.5 \ge b - 0.5$, and consequently $\max\{a - 0.5, 0\} \ge \max\{c - 0.5, 0\} \ge \max\{b - 0.5, 0\}$.

$$
\begin{aligned}
&|\max\{\max\{a,c\} - 0.5, 0\} - \max\{\max\{b,c\} - 0.5, 0\}| \\
&\quad = |\max\{a - 0.5, 0\} - \max\{c - 0.5, 0\}| \\
&\quad \le |\max\{a - 0.5, 0\} - \max\{b - 0.5, 0\}|. \\
&|\max\{\min\{a,c\} - 0.5, 0\} - \max\{\min\{b,c\} - 0.5, 0\}| \\
&\quad = |\max\{c - 0.5, 0\} - \max\{b - 0.5, 0\}| \\
&\quad \le |\max\{a - 0.5, 0\} - \max\{b - 0.5, 0\}|.
\end{aligned}
$$

5. $c \le a \le b$.

$$
\begin{aligned}
&|\max\{\max\{a,c\} - 0.5, 0\} - \max\{\max\{b,c\} - 0.5, 0\}| \\
&\quad = |\max\{a - 0.5, 0\} - \max\{b - 0.5, 0\}|. \\
&|\max\{\min\{a,c\} - 0.5, 0\} - \max\{\min\{b,c\} - 0.5, 0\}| \\
&\quad = |\max\{c - 0.5, 0\} - \max\{c - 0.5, 0\}| \\
&\quad = 0 \le |\max\{a - 0.5, 0\} - \max\{b - 0.5, 0\}|.
\end{aligned}
$$

6. $c \leq b \leq a$.

$$
\begin{aligned}
&|\max\{\max\{a,c\} - 0.5,0\} - \max\{\max\{b,c\} - 0.5,0\}| \\
&\quad = |\max\{a - 0.5,0\} - \max\{b - 0.5,0\}| . \\
&|\max\{\min\{a,c\} - 0.5,0\} - \max\{\min\{b,c\} - 0.5,0\}| \\
&\quad = |\max\{c - 0.5,0\} - \max\{c - 0.5,0\}| \\
&\quad = 0 \leq | \max\{a - 0.5,0\} - \max\{b - 0.5,0\}| .
\end{aligned}
$$

Thus, for every (a_1,a_2), (b_1,b_2), (c_1,c_2) T itholds that:

$$
\begin{aligned}
&|\max\{a_1 - 0.5,0\} - \max\{b_1 - 0.5,0\}| + \\
&|\max\{a_2 - 0.5,0\} - \max\{b_2 - 0.5,0\}| \geq \\
&\quad |\max\{\max\{a_1,c_1\} - 0.5,0\} - \max\{\max\{b_1,c_1\} - 0.5,0\}| + \\
&\quad |\max\{a_2 - 0.5,0\} - \max\{b_2 - 0.5,0\}| \geq \\
&\quad |\max\{\max\{a_1,c_1\} - 0.5,0\} - \max\{\max\{b_1,c_1\} - 0.5,0\}| + \\
&\quad |\max\{\min\{a_2,c_2\} - 0.5,0\} - \max\{\min\{b_2,c_2\} - 0.5,0\}| . \quad \blacksquare
\end{aligned}
$$

List of symbols

$(\Omega, \mathcal{A}, P)$	Probability space.
$X, Y, Z, \ldots$	Random variables.
$F_X, F_Y, F_Z, \ldots$	Cumulative distribution functions.
$(\Omega, \mathcal{A})$	Ordered space.
	Preference relation.
	Strict preference relation.
$\equiv$	Indifference relation.
	Incomparability relation.
$\mathcal{D}$	Set of random variables.
FSD	First degree stochastic dominance.
SSD	Second degree stochastic dominance.
nSD	n-th degree stochastic dominance.
$\mathcal{U}$	Set of increasing functions $u: \mathbb{R} \to \mathbb{R}$.
$\mathcal{U}$	Set of increasing and bounded functions $u: \mathbb{R} \to \mathbb{R}$.
Q	Probabilistic relation.
SP	Statistical preference.
$Q_{\frac{1}{2}}$	$\frac{1}{2}$-cut of the probabilistic relation Q.
C	Copula.
M	Minimum copula.
W	Łukasiewicz copula.
π	Product copula.
ϕ	Generator of an Archimedean copula.
$\mathcal{L}(\Omega)$	Set of gambles on Ω.
$\mathcal{K}$	Set of gambles $\mathcal{K}\ \mathcal{L}(\Omega)$.
$\underline{P}$	Lower prevision.

$\overline{P}$	Upper prevision.
$\mathcal{M}(\underline{P})$	Credal set associated with $\underline{P}$.
$\underline{E}$	Natural extension.
$\underline{F}$	Lower distribution function.
$\overline{F}$	Upper distribution function.
$(\underline{F}, \overline{F})$	P-box.
$\underline{P}_{(\underline{F},\overline{F})}$	Lower probability associated with the p-box $(\underline{F}, \overline{F})$.
Π	Possibility measure.
π	Possibility distribution.
N	Necessity measure.
$[\delta, \pi]$	Cloud.
Γ	Random set.
$\Gamma^*(A)$	Upper inverse of Γ in A.
$S(\Gamma)$	Set of measurable selections of the random set Γ.
P_{Γ}^*	Upper probability defined from the random set Γ.
$P_{*\Gamma}$	Lower probability defined from the random set Γ.
$P(\Gamma)$	Set of probabilities defined by the measurable selections.
$\beta_{[0,1]}$	Borel σ-algebra on $[0,1]$.
$\lambda_{[0,1]}$	Lebesgue measure on $[0,1]$.
$(C)\int f d\mu$	Choquet integral of f with respect to μ.
μ_A, ν_A	Membership and non-membership functions of an IF-set A.
π_A	Hesitation index of the IF-set A.
$[l_A, u_A]$	Lower and upper bounds of the IVF-set A.
$IFSs(\Omega)$	Set of all IF-sets defined on Ω.
$FS(\Omega)$	Set of all fuzzy sets defined on Ω.
$B(p)$	Bernoulli distribution with parameter p.
$Exp(\lambda)$	Exponential distribution with parameter λ.
$U(a, b)$	Uniform distribution in the interval (a, b).
$P_a(\lambda)$	Pareto distribution with parameter λ.
$\beta(p, q)$	Beta distribution with parameters p and q.
$\beta(p, q, a, b)$	Beta distribution on the interval (a, b) with parameters p and q.
$N(\mu, \sigma)$	Normal distribution with mean μ and variance σ^2.
$N(\mu, \Sigma)$	Multidimensional normal distribution with vector of means μ and matrix of variances-covariances Σ.

ρ	Correlation coefficient.
$Q_n(\ ,[\ ])$	Generalprobabilistic relation.
δ_a	Diracfunctional onthepoint a.
pp	Probabilisticprior relation.
sd	Strongdominance relation.
$X, Y, Z, \ldots$	Setsof randomvariables.
FSD i	Imprecise first degree stochastic dominance.
F_X, F_Y	Setsof cumulativedistribution functions.
bel	Belief function.
pl	Plausibility function.
SPi	Imprecise statistical preference.
$Q^{X,Y}$	Profile of preferences of the sets of random variables X and Y.
Ω_1, Ω_2	Ordered spaces.
$\overline{C}$	Upper copula.
$\underline{C}$	Lower copula.
$(\underline{C}, \overline{C})$	Imprecise copula.
$\underline{P}_X \quad \underline{P}_Y$	Independent natural extension of $\underline{P}_X$ and $\underline{P}_Y$.
$\underline{P}_X \quad \underline{P}_Y$	Strong product of $\underline{P}_X$ and $\underline{P}_Y$.
uo	Upper orthant relation.
lo	Lower orthant relation.
D_{IFS}	IF-divergence.
l_{IFS}	Hammingdistance forIF-sets.
d_H	Hausdorffdistance forIF-sets.
D_C, D_L	Hongand Kimdissimilarities.
D_O	Liet al. dissimilarity.
D_{HB}	Mitchell dissimilarity.
D_e^p, D_h^p	Liangand Shidissimilarities.
$D_{HY}^1, D_{HY}^2, D_{HY}^3$	Hungand Yang dissimilarities.
q_{IFS}	Euclideandistance forIF-sets.
D_s^p	Liangand Shidissimilarity.
D_C	Chen dissimilarity.
D_{DC}	Dengfenf and Chuntian dissimilarity.
$D_{\omega 1}, D_{pk1}, D_{pk3}$	Hungand Yang dissimilarities.
h_{IFS}	Function that defines a local IF-divergence.
$Cl([0,1])$	Setof cloudsdefinedon $[0,1]$
$IF\ ([0,1])$	Setof IF-setson $[0,1]$such that $\mu_A(x)=0$ and $\nu_A(y)=0$ for some x,y $[0,1]$

Index

List of Figures

List of Tables

Bibliography

[1] F. Aiche and D. Dub ois. An extension of sto chastic dominance to fuzzy random variables. In *Proceedings of IPMU 2010* . Springer-Verlag, 2010.

[2] M. A.Arcones, P. H.Kvam, and F.J. Samaniego. Nonparametric estimation ofa distribution sub ject to a stochastic preference constraint. *Journalof the American Statistical Association* , 97:170–182, 2002.

[3] B. Arnold. *Majorization andthe LorenzOrder: ABrief Introduction*. Springer, 1987.

[4] K. Atanas sov. Intuitionistic fuzzysets. In *Proceedings of VII ITKR, Sofia*,1983.

[5] K. Atanas sov. Intuitionistic fuzzysets. *Fuzzy Sets and Systems* , 20:87–96, 1986.

[6] K. Atanassov. Interval valued intuitionistic fuzzy sets. *Fuzzy SetsandSystems*, 31:343–349, 1989.

[7] K. Atanassov. *Intuitionistic Fuzzy Sets: Theory and Applications* . Physica-Verlang, Wyrzburg, 1999.

[8] K. Atanassov. Intuitionistic fuz zy sets: Past, present andfuture. In *Proccedings of the "European Society of Fuzzy Logic and Technology Conference" (EUSFLAT) 2003*, 2003.

[9] K. Atanassov and G. Gargov. Intuitionistic fuzzylogic. *CR Acad. Bulg. Soc.*, 43(3):9–12, 1990.

[10] K.AtanassovandG.Georgeiv. Intuitionisticfuzzyprolog. *Fuzzy Sets and Systems*, 53:121–128, 1993.

[11] A. Atkins on. On the measureme nt of p overty*Econometrica*, 55:749–764,1987.

[12] J. Au mann. Utility theory without the c ompleteness axiom. *Econometrica*, 30:445–462, 1962.

[13] J.Aumann. Integralofsetvaluedfunctions. *Journal of Mathematical Analysis and Applications*, 12:1–12, 1965.

[14] B. De Baets and H. De Meyer. Transitivity frameworks for recipro cal relations: cycle-transitivity versus -transitivity. *Fuzzy Sets and Systems*,152(2):249 – 270, 2005.

[15] B. De Baets, H. De Meyer, and K. De Lo of. On the cycle-transitivity of the mutual rank probability relation of a poset. *Fuzzy Sets and Systems* , 161(20):2695 – 2708, 2010.

[16] B. De Baets, H. De Meyer, B. De Schuymer, and S. Jenei. Cyclic evaluation of transitivity of recipro cal relations. *Social Choice and Welfare*, 26:217–238,2006.

[17] C. Baudrit and D. Dub ois. Practical representations of incomp lete probabilistic knowledge. *Computational Statistics and Data Analysis* , 51(1):86–108, 2006.

[18] F. Belzunce, J. A. Mercader, J. M. Ruiz, and F. Spizzichino. Sto chastic comparisons of multivariate mixture mo dels. *Journal of Multivariate Analysis*, 100:1657–1669, 2009.

[19] S. Benferhat, D. Dub ois,and H. Prade. Towards a possibilistic logic handling of preferences.*Applied Intel ligence*, 14:303–317, 2001.

[20] J. O.Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York, 1985.

[21] J. C. Bezdek, D. Spliman, and R. Spilman. A fuzzy relation space for groupdecision theory. *Fuzzy Sets and Syst ems*, 1:255–268, 1978.

[22] B. Bhandari, N. R. Pal, and D. D. Ma jumder. Fuzzy dive rge nce, probability measureof fuzzy events and image thresholding. *Pattern Recognition Letters* , 13:857–867, 1992.

[23] P. K. Bhatia and S. Singh. Anew measure of fuzzy dire cted divergence and its application in image segmentation. *Journal of Intel ligent Systems and Applications*, 4:81–89, 2013.

[24] P. Billingsley. *Probabilityand Measure*. Wiley-Inte rs cience, 3 edition, April 1995.

[25] P. J. Boland, M. Hollander, K. Joag-Dev, and S. Ko char. Bivariate dep endence prop erties of order statistics. *Journal of Multivariate Analysis* , 56:75–89, 1996.

[26] P. J. Boland, H. Singh, and B. Cukic. The sto chastic precedence ordering with applicationsin sampling and testing. *Journal of Applied Probability*, 41:73–82, 2004.

[27] B. Bouchon-M eunier, M. Rifqi, and S.Bothorel. Towards generalmeasuresof comparison of ob jects.*Fuzzy Sets and Systems* , 84:143–153, 1996.

[28] H. Brenner. Long-termsurvivalratesofcancer patientsachived bytheendofthe 20th century: a prio d analysis. *The Lancet* , 360:1131–1135, 2002.

[29] P. Burrillo an d H. Bus tince. Entropy onintuitionisticfuzzy setsand on interval-valued fuzzy sets. *Fuzzy Sets and Syst ems*, 78(3):305–316, 1996.

[30] H. Bus ti nce and P.Burrillo. Vague se ts are intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, 79:403–405,1996.

[31] T.Calvo, A.Kolesárová, M.Komorníková, andR.Mesiar. *Aggregation operators, New Trends and Applications*. Physica-Verlag, Heidelb ert, 2002.

[32] A. C as ero. Uso de simulación para el estudio de la preferencia estadística entre variables aleatorias. Traba jo Académicamente dirigido, Licenciatura de Matemáticas, Dirigido p or E. M irand a, 2011.

[33] D. Cayrac, D. Dub ois, M. Haziza, and H. Prade. Handling uncertainty with possibility theory andfuzzy sets in asatellite fault diagnosis application. *IEEE Transactions on Fuzzy Systems* , 4:251–259, 1996.

[34] T. Chaira and A. K. Ray. Segmentation using fuzzy divergence. *Pattern Recognit ion Letters*, 24:1837–1844, 2003.

[35] A. Chateauneuf, M. Cohen, and R. Kast. Comonotone random variable s in economics: a reviewofsome results. *Cahiers d'ECOMATH*,97.32,1997.

[36] S. M. Che n. Measures of similarity between vague sets. *FuzzySets andSystems*, 67:217–223, 1995.

[37] S.M. Chen. Sim ilarity measures b etween vague sets and b etween elements. *IEEE Trans. Syst. Mn Cybernet.* , 27(1):153–158, 1997.

[38] R.Cheng, B.Evans, andJ.Williams. Confidencebandestimationfordistributions used in probabilisticdesign. *International Journal of Mechanical Sciences* , 30:835–845, 1988.

[39] G. Cho quet. Theory of capacities. *Annales de l'Institut Fourier*,5:131–295,1953–1954.

[40] M.Condorcet. *An eassy on the application of probability theory to plurality decision making: an electionbetween three candidates* Reprintedin1989, F.Sommerlad, I. McLean (Eds.), 1785.

[41] O. Cordón, F. Herrera, F. Gomide, F. Hoffmann, and L. Magdalena. Ten years of genetic-fuzzy s ys temsa current frame work and new trends. In *Proccedings of Joint 9th IF SA World Congress and 20th NAFIPS International Conference, Vancouver-Canada* , pages 1241–1246. 2001.

[42] O. Cordón, F. Herrera, F. Hoffmann, and L. Magdalena. *Genetic Fuzzy Systems: Evolutionary Tuning and Learning of Fuzzy Knowledge Bases*. World Scientific Publishing, 2001.

[43] I. Couso and D. Dubois. An imprecise probability approach to joint extensions of stochastic and interval orderings. In *Proceedings of IPMU 2012*. Springer-Verlag, 2012.

[44] I. Couso, L. Garrido, S. Montes, and L. Sánchez. Expected pairwise comparison of the values of a fuzzy random variable. In C. Borgelt, G. González-Rodríguez, W. Trutschnig, M.A. Lubiano, M.A. Gil, P. Grzegiorzewski, and O. Hryniewicz, editors, *Combining Soft Computing and Statistical Methods in Data Analysis*, volume 77 of *Advances in Intelligent and Soft Computing*, pages 105–114. Springer, 2010.

[45] I. Couso, L. Garrido, and L. Sánchez. Similarity and dissimilarity measures between fuzzy sets: A formal relational study. *Information Sciences*, 229(0):122–141, 2013.

[46] I. Couso and L. Sánchez. The behavioural meaning of the median. In C. Borgelt, G. González-Rodríguez, W. Trutschnig, M.A. Lubiano, M.A. Gil, P. Grzegiorzewski, and O. Hryniewicz, editors, *Combining Soft Computing and Statistical Methods in Data Analysis*, volume 77 of *Advances in Intelligent and Soft Computing*, pages 115–122. Springer, 2010.

[47] Inés Couso, Serafín Moral, and Peter Walley. A survey of concepts of independence for imprecise probabilities. *Risk Decision and Policy*, 5:165–181, 2000.

[48] S.K. De, R. Biswas, and A.R. Roy. An application of intuitionistic fuzzy sets in medical diagnosis. *Fuzzy Sets and Systems*, 117:209–213, 2001.

[49] B. De Baets and H. De Meyer. On the cycle-transitive comparison of artificially coupled random variables. *International Journal of Approximate Reasoning*, 47:306–322, 2008.

[50] G. de Cooman. Integration in possibility theory. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals – Theory and Applications*, pages 124–160. Physica-Verlag (Springer), Heidelberg, 2000.

[51] G. de Cooman and D. Aeyels. Supremum preserving upper probabilities. *Information Sciences*, 118:173–212, 1999.

[52] G. de Cooman, E. Miranda, and M. Zaffalon. Independent natural extension. *Artificial*, 175(12–13):1911–1950, 2011.

[53] G. de Cooman, M. C. M. Troffaes, and E. Miranda. n-Monotone exact functionals. *Journal of Mathematical Analysis and Applications*, 347:143–156, 2008.

[54] H.DeMeyer, B.DeBaets, andB.DeSchuymer. Onthetransitivityofthecomonotonic and cou ntermonotonic comparison of random variables. *Journal of Multivariate Analysis* , 98:177–193, 2007.

[55] B. De Schuymer, H. De Meyer, and B. De Baets. *A fuzzy approach to st ochastic dominance of random variables*, pages 253–260. Lecture Notes in Artifial Intelligence 2715, 2003.

[56] B. De Schuymer, H. De Meyer, an d B. De Baets. Cycle-transitivitycomparison of indep endent random variables. *Journal of Multivariate Analysis*, 96:352–373, 2005.

[57] B. De Schuymer, H. De Meyer, B. De Baets, and S. Jenei. On the cycle-transitivity of the dice mo del. *Theory and Decision* , 54:261–285, 2003.

[58] B. De Schuymer, H. De Me yer, and B. De Baets. Extremecopulas and thecomparison of ordered lists. *Theory and Decision* , 62:195–217, 2007.

[59] A. P.Dempster. Upp er and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics* , 38:325–339, 1967.

[60] D. Denneb erg. *Non-Additive Measure and Int egral*. KluwerAcademic, Dordrecht, 1994.

[61] T. Deno eux. Extending sto chastic ordering to b elief func tions on the real line. *Information Sciences* , 179:1362–1376, 2009.

[62] M. Denuit, J. Dhaene, M. Goovaerts, and R. Kaas. *Actuarial Theory for Dependent Risks Measures, Orders and Models* John Wiley& Sons,2005.

[63] G. Deschrijverand E.Kerre. A generalization of op erators on intuitionistic fuzzy sets using triangular norms and c on orms. *Notes on Intuitionistic Fuzzy Sets*,1:19–27, 2002.

[64] S. Destercke, D. Dub ois, and E. Cho jnacki. Unifying practical uncertainty representations- ii: Generalised p-b oxes. *International Journal of Approximate Reasoning*, 49(3):649–663, 2008.

[65] S. Destercke, D. Dub ois, and E. Cho jnacki. Unifying practical uncertainty representations: Ii.clouds. *International Journal of Approximate Reasoning*,49(3):664–677, 2008.

[66] P. Dirac. *Principles of quantu m mechanics*. Oxford attheClaredon Press, 4th edition, 1958.

[67] D. Dub ois, H. Fargier, and P. Perny. Qualitative decision theory withpreference relations and comparative uncertainty: An axiomaticapproach. *Artificial Intel ligence*, 148:219–260, 2003.

[68] D. Dubois, S. Gottwald, P. Hajek, J. Kacprzyk, and H. Prade. Terminological difficulties in fuzzy set theory-the case of "intuitionistic fuzzy sets". *Fuzzy Sets and Systems*, 156:485–491, 2005.

[69] D. Dubois, E. Kerre, and R. Mesiar. Fuzzy interval analysis. In D. Dubois and H. Prade, editors, *Fundamentals of Fuzzy sets*, pages 483–581. Kluwer Academic Publishers, Boston, 2000.

[70] D. Dubois and H. Prade. *Possibility Theory*. Plenum Press, New York, 1988.

[71] D. Dubois and H. Prade. *Fundamentals of Fuzzy Sets*. Kluwer Academic Publishers, Massachusetts, 2000.

[72] D. Dubois and H. Prade. An overview of the asymmetric bipolar representation of positive and negative information in possibility theory. *Fuzzy Sets and Systems*, 160(10):1355–1366, 2009.

[73] D. Dubois and H. Prade. Gradualness, uncertainty and bipolarity: Making sense of fuzzy sets. *Fuzzy Sets and Systems*, 192:3–24, 2012.

[74] J. Fan and W. Xie. Distance measure and induced fuzzy entropy. *Fuzzy Sets and Systems*, 104:305–314, 1999.

[75] S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers, and K. Sentz. Constructing probability boxes and Dempster-Shafer structures. Technical Report SAND2002–4015, Sandia National Laboratories, January 2003.

[76] S. Ferson and W. Tucker. Probability boxes as info-gap models. In *Proceedings of the Annual Meeting of the North American Fuzzy Information Processing Society*, New York (USA), 2008.

[77] M. Finkelstein. On stochastic comparisons of population densities and life expectancies. *Demographic Research*, 13:143–162, 2005.

[78] P. C. Fishburn. *Interval Orders and Interval Graphs*. Wiley, New York, 1985.

[79] M. Fréchet. Généralisations du théorème des probabilités totales. *Fundamenta Mathematicae*, 25:379–387, 1935.

[80] J. L. García-Lapresta and B. Llamazares. Majority decisions based on differences of votes. *Journal of Mathematical Economics*, 35:463–381, 2001.

[81] I. Georgescu. *Fuzzy Choice Functions, a Revealed Preference Approach*. Springer, Berlin, 2007.

[82] I. Gilboa and D. Schmeidler. Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, 18:141–153, 1989.

[83] F. J. Giron and S. Rios. Quasi-Bayesian behaviour: A more realistic approach to decision making? In J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith, editors, *Bayesian Statistics*, pages 17–38. Valencia University Press, Valencia, 1980.

[84] I. R. Goodman. Fuzzy sets as equivalence classes of random sets. In R.R. Yager, editor, *Fuzzy Sets and Possibility Theory*, pages 327–342. Pergamon Press, Oxford, 1982.

[85] P. Grzegorzewski. Distances between intuitionistic fuzzy sets and/or interval-valued fuzzy sets based on the hausdorff metric. *Fuzzy Sets and Systems*, 148:319–328, 2004.

[86] P. Grzegorzewski and E. Mrowka. Probability of intuitionistic fuzzy events. In P. Grzegiorzewski, O. Hryniewicz, and M. A. Gil, editors, *Soft methods in probability, statistics and data analysis*, pages 105–115. Physica-Verlag, 2002.

[87] F. Herrera. Genetic fuzzy systems: taxonomy and current research trend and prospects. *Evolutionary Intelligence*, 1:27–46, 2008.

[88] C. J. Himmelberg. Measurable relations. *Fundamenta Mathematicae*, 87:53–72, 1975.

[89] D. H. Hong and C. Kim. A note on similarity measures between vague sets and between elements. *Information Sciences*, 115:83–96, 1999.

[90] T. Hu and C. Xie. Negative dependence in the balls and binds experiment with applications to order statistics. *Journal of Multivariate Analysis*, 97:1342–1354, 2006.

[91] E. Hüllermeier. *Cased-based approximate reasoning*. Springer, 2007.

[92] W-L. Hung and M-S. Yang. Similarity measures of intuitionistic fuzzy sets based on hausdorff distances. *Pattern Recognition Letters*, 25:1603–1611, 2004.

[93] W-L. Hung and M-S. Yang. Similarity measures of intuitionistic fuzzy sets based on l_p metric. *International Journal of Approximate Reasoning*, 46:120–136, 2007.

[94] W-L. Hung and M-S. Yang. On similarity measures between intuitionistic fuzzy sets. *Internation Journal of Intelligent Systems*, 23:364–383, 2008.

[95] M. Islam and J. Braden. Bio-economics development of floodplains: Farming versus fishing in bangladesh. *Environment and Development Economics*, 11(1):95–126, 2006.

[96] J.-Y. Jaffray and M. Jeleva. Information processing under imprecise risk with the Hurwicz criterion. In *Proccedings of the Fifth Int. Symposium on Imprecise Probabilities and Their Applications, ISIPTA*, 2007.

[97] V. Janiš. T-norm based cuts of intuitionistic fuzzy sets. *Information Sciences*, 180(7):1134–1137, 2010.

[98] A. Kaur, L. S. Prakasa Rao, and H. Singh. Testing for second-order sto chastic dominanceof two distributions. *Econometric Theory* , 10:849–866, 1994.

[99] E. P. Klement, R. Mesiar, and E. Pap. *Triangular norms*. Kluwer Academic Publishers, Dordrecht, 2000.

[100] E. P. Klement, R. Mesiar, and E. Pap. *Triangular Norms*. Kluwer Academic Publishers, Dordrecht, 2000.

[101] G. Klir and B.Yuan. *Fuzzy setsandfuzzy logic: theory andapplications*. Prentice Hall, New Jersey, 1995.

[102] V. Krätschmer. When fuzzy measures are upp er envelop es of probability measures. *Fuzzy Sets and Syst ems*, 138:455–468, 2003.

[103] E. Kriegler. Utilizing belief functions for the estimation of future climate change. *International Journal of Approximate Reasoning*, 39(2–3):185–209,2005.

[104] R. Kruse and K. D. Meyer. *Statisticswithvague data*. D. Reidel Publishing Company, Dordrecht, 1987.

[105] H.Kwakernaak. Fuzzyrandomvariables, i, ii. *Information Sciences*, 15-17:253–178, 1979.

[106] E.L. Lehmann. Orderingfamilies ofdistributions. *Annals of Mathematical Statistics* , 26:399–419, 1955.

[107] I. Levi. *The enterprise of know ledge*.MIT Press,Cambridge, 1980.

[108] H. Levy. Stochastic dominance and exp ected utility: survey andanalysis. *Management Science* , 38:555–593, 1992.

[109] H. Levy. *Stochastic Dominance*.KluwerAcademic Publishers,1998.

[110] M.Levy andH.Levy. Testin g for risk aversion:a sto chastic dominance approach. *Economics Letters* , 71:233–240, 2001.

[111] D. F. Li andC.T. Chen. New similarity measures ofintuitionisticfuzzy sets and applications to pattern recognitions. *Pattern Recognition Letters* , 23:221–225, 2002.

[112] X.Liand X.Hu. Some new sto chastic comparisons for redundancy allo cations in seriesand parallel systems. *Statistics and Probability Letters*,78:3388–3394,2008.

[113] Y. Li, D. L. Olson, and Z. Qin. Similarity measures between intuitionistic fuzzy (vague) sets: Acomparative analysis. *Pattern RecognitionLetters*, 28:278–285, 2007.

[114] Z. Liang and P. Shi. Similarity measures on intuitionistic fuzzy sets. *Pattern Recognition Letters*, 24:2687–2693, 2003.

[115] C.J. Liau. On the possibility theory-based semantics for logics of preference. *International Journal of Approximate Reasoning*, 20:173–190, 1999.

[116] P. Limbourg. Multiobjective optimization of problems with epistemic uncertainty. In *Proceedings of Third International Conference on Evolutionary Multi-Criterion Optimization, Guanajuato-Mexico*, pages 413–427, 2005.

[117] M. Lorenz. Methods of measuring the concrentration of wealth. *Publications of the American Statistical Association*, 9(209–219):70, 1905.

[118] P. Lui, F. Jin, X. Zhang, Y. Su, and M.Wang. Research on the multi-attribute decision-making under risk with interval probability based on prospects theory and the uncertain linguistic variables. *Knowledge-Based Systems*, 24:554–561, 2011.

[119] X. Lui. Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets and Systems*, 52:305–318, 1992.

[120] D. Martinetti, V. Janiš, and S. Montes. Cuts of intuitionistic fuzzy sets respecting fuzzy connectives. *Information Sciences*, 232:267–275, 2013.

[121] D. Martinetti, I. Montes, and S. Díaz. Transitivity of the fuzzy and the probabilistic relations associated to a set of random variables. In P. Burrillo, H. Bustince, B. De Baets, and J. Fodor, editors, *Proceedings of EUROFUSE 2009 Conference*, pages 21–26, 2009.

[122] D. Martinetti, I. Montes, S. Díaz, and S. Montes. A study of the transitivity of the probabilistic and fuzzy relation. *Fuzzy Sets and Systems*, 184:156–170, 2011.

[123] J. M. Merigó and A. M. Gil-Lafuente. Induced 2-tuple linguistic generalized aggregation operators and their application in decision-making. *Information Sciences*, 236:1–16, 2013.

[124] P. Mikusiński, H. Sherwook, and M.D. Taylor. The Fréchet bounds revisited. *Real Anal Exchange*, 17:759–764, 1991-92.

[125] E. Miranda. *Análisis de la información probabilística de los conjuntos aleatorios*. PhD thesis, University of Oviedo, 2003. In Spanish.

[126] E. Miranda. A survey of the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 48(2):628–658, 2008.

[127] E. Miranda, I. Couso, and P. Gil. Upper probabilities and selectors of random sets. In P. Grzegorzewski, O. Hryniewicz, and M.A. Gil, editors, *Soft methods in probability, statistics and data analysis*, pages 126–133. Physica-Verlag, Heidelberg, 2002.

[128] E. Miranda, I. Couso, and P. Gil. A random set characterization of possibility measures. *Information Sciences*, 168:51–75, 2004.

[129] E. Miranda, I. Couso, and P. Gil. Random intervals as a model for imprecise information. *Fuzzy Sets and Systems*, 154:386–412, 2005.

[130] E. Miranda, I. Couso, and P. Gil. Approximations of upper and lower probabilities by measurable selections. *Information Sciences*, 180:1407–1417, 2010.

[131] E. Miranda, I. Couso, and P. Gil. Upper probabilities attainable by distributions of measurable selections. *Information Sciences*, 180(8):107–1417, 2010.

[132] E. Miranda, G. de Cooman, and I. Couso. Lower previsions induced by multi-valued mappings. *Journal of Statistical Planning and Inference*, 133(1):173–197, 2005.

[133] E. Miranda, G. de Cooman, and E. Quaeghebeur. Finitely additive extensions of distribution functions and moment sequences: The coherent lower prevision approach. *International Journal of Approximate Reasoning*, 48(1):132–155, 2008.

[134] E. Miranda and I. Montes. Imprecise preferences by means of probability boxes. In *Proceedings of 4th ERCIM*, page 63, 2011.

[135] E. Miranda, I. Montes, R. Pelessoni, and P. Vicig. Bivariate p-boxes. In *Proceedings of 8th ISIPTA*, 2013.

[136] E. Miranda, M. Troffaes, and S. Destercke. Generalised p-boxes on totally ordered spaces. In D. Dubois, M. Lubiano, H. Prade, M. Gil, P. Grzegiorzewski, and O. Hryniewicz, editors, *Soft Methods for Handling Variability and Imprecision*, pages 235–242. Springer, 2008.

[137] E. Miranda and M. Zaffalon. Coherence graphs. *Artificial Intelligence*, 173(1):104–144, 2009.

[138] H. B. Mitchell. On the dengfeng–chuntian similarity measure and its application to pattern recognition. *Pattern Recognition Letters*, 24:3101–3104, 2003.

[139] A. Müller and D. Stoyan. *Comparison Methods for Stochastic Models and Risks*. Wiley, 2002.

[140] I. Montes. Some general comments about statistical preference. In *Proceedings of I TUIM Conference*, 2010.

[141] I. Montes. Divergences for intuitionistic fuzzy sets. In *Proceedings of II TUIM Conference*, 2011.

[142] I. Montes. Comparison of more than two random variables by means of the statistical preference. In *Proceedings of XXXIII SEIO Conference*, 2012.

[143] I. Montes, S. Díaz, and S. Montes. Comparison of imprecise fitness values modelled by beta distributions. pages 489–494, 2010cited By (since 1996)0.

[144] I.Montes, J.Hernández, D.Martinetti, andS.Montes. Characterizationofcontinuous t-normscompatible withzadeh's probability of fuzzy events. *Fuzzy Setsand Systems*, 228:29–43, 2013.

[145] I. Montes, V. Janiš, and S. Montes. Anaxiomatic definitionof divergence for intuitionistic fuzzy sets. volume 1,pages547–553,2011. citedBy(since 1996)0.

[146] I.Montes, V.Janiš, and S.Montes. Onthe study of some measuresof comparison of if-sets. In *Proceedings of XVI ESTYLF Conference*, 2011.

[147] I. Montes, V. Janiš, and S. Montes. Local if-divergences. *Communications in Computer and Information Science*, 298CCIS(PART 2):491–500, 2012. cited By (since 1996)0.

[148] I.Montes, D.Martinetti, S.Díaz, andS.Montes. Acharacterisationof statistical preferenceby means of the median. 2012. Submitted forpublication.

[149] I. Montes, D. Martinetti, and S. Díaz. On the statistical preference as a pairwise comparison for random variables. In *Proceedings of EUROFUSE 2009 Conference*, 2009.

[150] I. Montes, D. Martinetti, S. Díaz, and S. Montes. Comparisonof random variables coupled byarchimedean copulas. In Christian Borgelt, Gil González-Rodríguez, Wolfgang Trutschnig, María Lubiano, María Gil, Przemyslaw Grzegorzewski, and OlgierdHryniewicz, editors, *Combining Soft Computing and Statistical Methods in Data Analysis*, volume 77of *Advances in Intelligent and Soft Computing*, pages 467–474. Springer Berlin / Heidelberg, 2010.

[151] I. Montes, D. Martinetti, S. Díaz, and S. Montes. Estudio de la preferencia estadísticaen distribuciones normales bidimensionales. In *Proceedings of XXXII SEIO Conference*, 2010.

[152] I.Montes, D.Martinetti, S.Díaz, andS.Montes. Statistical preference asacomparison method of two imprecise fitness values. In *Proceedings of XV ESTYLF Conference*, 2010.

[153] I. Montes, D. Martinetti, S. Díaz, and S. Montes. Statistical preference as a tool in consensus processes. In Enrique Herrera-Viedma, José García-Lapresta, Janusz Kacprzyk, Mario Fedrizzi, Hannu Nurmi, and SlawomirZadrozny, editors, *Consensual Processes*, volume 267 of *Studies in Fuzziness and Soft Computing*, pages 65–92. Springer Berlin / Heidelberg, 2011.

[154] I. Montes, D. Martinetti, and S. Montes. Statistical preference for testing algorithms. In *Proceedings of Centennial Congress of the Spanish Royal Mathematical Society RSME*, 2011.

[155] I. Montes, E. Miranda, and S. Díaz. Preferences with p-boxes. In *Proceedings of 4th WPMSIIP*, 2011.

[156] I. Montes, E. Miranda, and S. Montes. Decision making with imprecise probabilities and utilities by means of statistical preference and stochastic dominance. *European Journal of Operational Research*, 234(1):209–220, 2014.

[157] I. Montes, E. Miranda, and S. Montes. Stochastic dominance with imprecise information. *Computational Statistics and Data Analysis*, 71/C:867–885, 2014. In press.

[158] I. Montes, N.R. Pal, V. Janiš, and S. Montes. Divergence measures for intuitionistic fuzzy sets. *IEEE Transactions on Fuzzy Systems*, Accepted to publication, 2014.

[159] S. Montes. *Partitions and divergence measures in fuzzy models*. PhD thesis, University of Oviedo, 1998.

[160] S. Montes, I. Couso, P. Gil, and C. Bertoluzza. Divergence measure between fuzzy sets. *International Journal of Approximate Reasoning*, 30:91–105, 2002.

[161] S. Montes, T. Iglesias, V. Janiš, and I. Montes. A common framework for some comparison measures of if-sets. In *Proceedings of IWIFSGN Conference*, 2012.

[162] S. Montes, D. Martinetti, I. Montes, and S. Díaz. Graded comparison of imprecise fitness values. In *Proceedings of Hard & Soft Conference*, 2010.

[163] S. Montes, D. Martinetti, I. Montes, and S. Díaz. Min-transitivity of graded comparisons for random variables. In *Proceedings of FUZZ-IEEE 2010, Barcelona-Spain*. 2010.

[164] R. F. Nau. Indeterminate probabilities on finite sets. *The Annals of Statistics*, 20:1737–1767, 1992.

[165] R. F. Nau. The shape of incomplete preferences. *Annals of Statistics*, 34(5):2430–2448, 2006.

[166] R. Nelsen. *An introduction to copulas*. Springer, New York, 1999.

[167] R. Nelsen, J. J. Quesada Molina, J. A. Rodríguez-Lallena, and M. Úbeda Flores. Best-possible bounds on sets of bivariate distribution functions. *Journal of Multivariate Analysis*, 90:348–358, 2004.

[168] A. Neumaier. Clouds, fuzzy sets and probability intervals. *Reliable Computing*, 10:249–272, 2004.

[169] H. T. Nguyen. On random sets and belief functions. *Journal of Mathematical Analysis and Applications*, 65(3):531–542, 1978.

[170] H.T. Nguyen. *An introduction to random sets*. Chapman and Hall, 2006.

[171] N. Noyan, G. Rudolf, and A. Ruszczyński. Relaxations of linear programming problems with first order sto chastic dominance constrains. *Operations ResearchL*, 34:653–659, 2006.

[172] M. Ob erguggenb erger and W.Fellin. Reliability bounds through random sets: non-parametric metho ds and geotechnical applications. *Computers and Structures*, 86(10):1093–1101, 2008.

[173] W. Ogryczak and A.Ruszczyński. From sto chastic dominance to mean-risk mo d-els: Semideviations and riskmeasures. *European Journal of Operational Research*, 116:33–50, 1999.

[174] G. Owe n. *Game theory*. Academic Press, San Die go. Third Edition, 1995.

[175] N. R. Pal and S. K. Pal. Entropy, a new definitionand its applications. *IEEE Trans. Syst. Mn Cybernet.* , 21(5):1260–1270, 1991.

[176] R.Pelessoni, P.Vicig, I.Montes, andE.Miranda. Imprecise copulasand bivariate sto chastic orders.In *Proceedings of EUROFUSE 2013 Conference*,2013.

[177] E. Raufas te,R. DaSilva Neves, and C. Mariné. Testingthe descriptive validity of possibility theory in human judgements of uncertainty. *Artificial Intel ligence*, 148:197–218, 2003.

[178] D. Ríos Insua. On the foundations of decision analysis with partial information. *Theory and Decision* , 33:83–100, 1992.

[179] M. Roub ens and Ph. Vincke. *Preference Model ling*.Springer-Verlag,Berlin, 1985.

[180] L. Sánchez an d I. CousoAdvo cating the use of imprecise observed data in genetic fuzzy systems. *IEEE Transactions on Fuzzy Systems* , 15:551–562, 2007.

[181] L. Sánchez, I. Couso, and J. Cas illas.A multiob jective genetic fuzzy system with imprecise probability fitness forvague data. In *Proceedings of 2006 IEEE International Conference on Evolutionary Fuzzy Sy st ems, Ambleside-UK*. 2006.

[182] L. Sánchez,I. Couso, and J. Casillas. Mo deling vague data with genetic fuzzy system under a combination of crisp and imprecise criteria. In *Proceedings of 2007 IEEE Symposionon Computational Intel ligence in Multicriteria Decision Making, Honolulu-USA*. 2007.

[183] L. Sánchez, I. Couso, and J. Casil las.Geneticlearningof fuzzyrulesbased onlow quality data. *Fuzzy Sets and Systms* , 160:2524–2552, 2009.

[184] J. K.Satia andR. E. Lave. Markovian decision pro cesses with uncertain transition probabilities. *Operations Research*,21:728–740,1973.

[185] M. Scarsini. Copulae of capacities on pro duct spaces. In LudgerRüschendorf, Berthold Schwe izer, and Michael Dee Taylor, editors, *Distributions withFized Marginals andRelated Topics*, volume28 of *IMS Lecture Notes- Monograph Series*, pages 307–318. Institute ofMathematical Statistics, 1996.

[186] T. Seid enfeld,M. J. Schervish, and J. B. Kadane. Arepresentation of partially ordered preferences.*The Annals of Statistics* , 23:2168–2217, 1995.

[187] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.

[188] M. Shaked and J. G. Shanthikumar. *Stochastic Orders and theirapplications*. Springer, 2006.

[189] A. Sklar. Fonctions de répartitionà n-dimensions et leursmarges. *Publications de l'Institute de Statist iqu e de l'Université de Paris*, 8:229–231, 1959.

[190] C.Strobl, A.L.Boulesteix, andT.Augustin. Unbiased split selection for class ification trees based on th e gini index. *Computational Statistics and Dat a Analysis*, 52(1):483–501, 2007.

[191] Surveillance, Epidemiology, and EndResults(SEER)ProgramPublic-UseData (1973-1998). *Bethesda (MD):National Cancer Institute, DCCPS,Cancer Surveillance Reasearch Program, Cancer StatisticsBranch, released April2001, based on the August 2000 submission*.

[192] R.Szekli. *Stochastic ordering and dependence in applied probability*. Springer,1995.

[193] E. Szmidt and J.Kacprzyk. Distances between intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, 114:505–518, 2000.

[194] E. Szmidt and J. Kacprzyk. An applicationof intuitionistic fuzzyset similarity measures to a multi-criteria decision makingproblem. InLeszek Rutkowski, Ryszard Tadeusiewicz, LotfiA. Zadeh, and JacekM. Å»urada, editors, *Artificial Intel ligence and Soft Computing ICAISC 2006*, volume 4029 of *Lecture Notes in Computer Science*, pages 314–323. Springer Berlin Heidelb erg, 2006.

[195] E. Szmidtand J.Kacprzyk. Analysis of sim ilarity measures for atannasov's intuitionistic fuzzy sets. In J. P. Carvalho, D. Dub ois, U. Kaymak, and J. M. C. Sousa, editors, *Proccedings of"International Fuzzy Systems Association World Congress and 2009 EuropeanSociety of Fu zzy Logic and Technology Conference"(IFSA-EUSFLAT 2009)*,pages 1416–1421. 2009.

[196] H. Tanaka and P. J. Guo. *Possibilistic Data Analysis forOperations Research*. Physica-Verlag, Heidelb erg, 1999.

[197] J. Teich. Pareto-front exploration with uncertain objectives. In *Proceedings of First International Conference on Evolutionary Multi-Criterion Optimization*. 2001.

[198] M. Troffaes and S. Destercke. Probability b oxes on totally preordered spaces for multivariate mo dell. *International Jou rnal of Approximate Reasoning*, 52(6):767–791, 2011.

[199] M. Troffaes, E. Miranda, an d S. Destercke. On the connection b etween probabil ity b oxes and p ossibility measures. In *Proceedings of EUSFLAT'2011*, pages 836–843, Aix-les-Bains, France, 2011.

[200] M. Troffaes, E. Miranda, an d S. Destercke. On the connection b etween probabil ity b oxes and possibility measures. *Information Sciences*, 224:88–108, 2013.

[201] M. C.M. Troffaes. *Optimality, Uncertainty, and Dynamic Programming with Lower Previsions*. PhD th esis, Ghent University, Ghent, Belgium, March 2005.

[202] M.C. M.Troffaes. Decision makingunder uncertainty usingimprecise probabilities. *International Journal of Approximate Reasoning*, 45(1):17–29,2007.

[203] P. Wakker. Characterizing optimism and p essimism through comonoton icity. *Journal of Economic Theory*, 52:453–463,1990.

[204] P. Walley. Coherent lower (and upp er) probabilities. Technical Rep ort Statistics Research Rep ort 22, University of Warwick, Coventry, 1981.

[205] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman andHall, London, 1991.

[206] W. Wang and X. Xin. Distan ce measure b etween intuitionistic fuzzy sets. *Pattern Recognit ion Letters*, 26:2063–2069, 2005.

[207] X. Wang, C. Wu, and X.Wu. Choice functionsinfuzzy environment: An overview. *Studies in Fuzziness and Soft Compu ting*, 261:149–169,2010. citedBy (since 1996)2.

[208] G. A. Whitmore and M. C. Findlay. *Stochastic dominance: An approachto decision-making under risk*. Lexington Bo oks, 1978.

[209] R. Williamson. *Probabilistic arithmetic*. PhD thesis, Universi ty of Queensland, 1989.

[210] C. WratherandP.L.Yu. Probabilitydominance inrandomoutcomes. *Journal of Optimization Theory and Applications*, 36:315–334,1982.

[211] Z. Xu. Some similarity measures of intuiti on istic fuzzy sets and their applications to multiple attribute decisionmaking. *Fuzzy Optimizat ion and Decision Making*, 6:109–121, 2007.

[212] Z. S. Xu and J. Chen. An overview of distance an d similarity measures of intuitionistic fuzzy sets. *Internation Journal of Uncertainty, Fuzziness and Know ledge-Based Systems*, 16(4):529–555, 2008.

[213] Y. Yuan and H. Zhuang. A gen eral algorithm for generating fuzzy classification rules. *Fuzzy Sets and Syst ems*, 84:1–19, 1996.

[214] L. A. Zad eh. Fuzzy sets. *Information and Control* , 8:338–353, 1965.

[215] L. A.Zadeh. Probability measuresof fuzzy events. *Journal of Mathematical Analysis and Applications* , 23:421–427, 1968.

[216] L. A. Zadeh. Outline ofa newapproach tothe analysisofcomplex systemsand decision pro cesses interval-valued fuzzy sets. *IEEE Trans. Syst. Mn Cybernet.*, 3:28–44, 1973.

[217] L. A. Zadeh. Fuzzy sets as a basis fora theory of possibility. *Fuzzy Setsand Systems*, 1:3–28, 1978.

[218] M. Zaffalon and E. Miranda. Conservativeinference ruleforuncertain reasoning under incompleteness. *Journal of Artificial Intel ligence Research*, 34:757–821, Jan 2009.

[219] M.Zaffalon, K.Wesnes, andO.Petrini. Reliablediagnosesofdementiabythenaive credal classifier inferred from incomplete cognitive data. *Artificial Intel ligence in Medicine*, 29:61–79, 2003.